KU LEUVEN
GROUP BIOMEDICAL SCIENCES
FACULTY OF MEDICINE
DEPARTMENT OF DEVELOPMENT AND REGENERATION
LABORATORY OF MOLECULAR BIOLOGY (CELGEN)

The role of Zeb2 in cell fate decisions during development

Agata Stryjewska

Promoter:	Prof. Danny Huylebroeck
Co-promoter:	Prof. Eve Seuntjens
Chair:	Prof. Maurilio Sampaolesi
Examining committee:	Prof. Susana Chuva de Sousa Lopes
	Prof. Luc Leyns
	Prof. Frank Luyten
	Prof. Derk ten Berge
	Prof. An Zwijsen

Doctoral Thesis In Biomedical Sciences
Leuven, 2016

Copyright 2016 KU Leuven, Faculty of Medicine, Biomedical Sciences
Uitgegeven in eigen beheer, Agata Stryjewska
Herestraat 49 – box 700, B-3000 Leuven, Belgium

Alle rechten voorbehouden. Niets uit deze uitgave mag worden vermenigvuldigd en/of openbaar gemaakt worden door middel van druk, fotokopie, microfilm, elektronisch of op welke andere wijze ook zonder voorafgaande schriftelijke toestemming van de uitgever.

All rights reserved. No part of the publication may be reproduced in any form of print, photoprint, microfilm, electronic or any other means without written permission from the publisher.

Acknowledgments

To my colleagues, collaborators and teachers.

I would like to thank everyone who helped me during my PhD.

Inspiration, motivation and knowledge are essential to pursue a career in science.

For me the **inspiration** comes from interacting with researchers who are truly interested in and passionate about science. I very much enjoyed working with people who are driven and whose way of thinking is not limited by what is already known and well established. Thank you to everyone who inspired and encouraged me to stay open-minded and told me to never stop asking questions and to believe in myself.

Everyone knows that failure is a part of success but, in real life, it is not always easy to stay motivated and keep working on achieving your goals. People I consider the most successful, both in science and in life, have taught me a lot about **motivation** and determination. I learned from them how to deal with failures and remain focused on the end goal. Thank you for being a great example of how to conquer adversity without losing enthusiasm.

As a scientist you never stop learning. I truly appreciate all the teachers and experts who dedicated their time and shared their **knowledge** with me. You are a part of my success.

Last, I would like to thank every kind, humble and honest person I crossed paths with when working in the lab. It was a pleasure.

Once again a very big thank you to everyone who helped me along the way.

Agata

Table of contents

List of abbreviations

2i	Two inhibitors, ground-state medium for ESC culture
aa	Amino acid
AGM	Aorta – gonadal – mesonephros
ANR	Anterior neural ridge
Bmp	Bone morphogenetic proteins
bp	Base pair
cDNA	Complementary DNA
CFU	Colony-forming unit
CGE	Caudal ganglionic eminence
CGI	CpG island
ChIP	Chromatin immunoprecipitation
ChIP-PET	Chromatin immunoprecipitation-Pair-end tags
CID	CtBP interaction domain
CNS	Central nervous system
CpG	Cytosine followed by a guanine nucleotide in the linear sequence of bases along its 5' → 3' direction
CtBP	C-terminal binding protein 1
CZF	C-terminal zinc finger cluster
d	Day (of differentiation)
DEGs	Differentially expressed genes
DMR	Differentially methylated region
DNA	Deoxyribonucleic acid
E	Embryonic day
EB	Embryoid body
EGC	Embryonic germ cells
EMT	Epithelial-to-mesenchymal transition
Epi	Epiblast
EpiSC	Epiblast stem cells
ESC	Embryonic stem cell
FC	Fold change
GOF	Gain of function
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
HSC	Hematopoietic stem cell
ICM	Inner cell mass
IHC	Immunohistochemistry
IN	Interneuron
iPSC	Induced pluripotent stem cell
ISH	*In situ* hybridization
KD	Knockdown
KO	Knockout
LGE	Lateral ganglionic eminence

LIF	Leukemia inhibitory factor
LincRNA	Large intergenic RNA
LncRNA	Long non-coding RNA
LOF	Loss of function
MET	Mesenchymal-to-epithelial transition
MGE	Medial ganglionic eminence
MHB	Midbrain-hindbrain border
miRNA	Micro RNA
Mut	Mutant
MOWS	Mowat-Wilson syndrome
nLacZ	Nuclear lactose operon encoding Beta-galactosidase
NuRD	Nucleosome remodeling and deacetylation complex
NZF	N-terminal zinc finger cluster
PCA	Principal component analysis
PcG	Polycomb group proteins
PE	Primitive endoderm
POA	Preoptic area
R26	Rosa 26 locus
REST	Repressor element 1 (RE-1) silencing transcription factor
RNA	Ribonucleic acid
RNA-Seq	RNA sequencing
RRBS	Reduced representation bisulfate sequencing
SBD	Smad-binding domain
SFEB	Serum-free suspension culture
shRNA	Short hairpin RNA
Sip1	Smad-interacting protein 1
SVZ	Subventricular zone
TAD	Topologically associated domain
TE	Trophoectoderm
Tet1	Tet methylcytosine dioxygenase 1
Tgfβ	Transforming growth factor beta
TPM	Transcripts per million (reads)
TSS	Transcription start site
VT	Ventral telencephalon
Wnt	Stands for "Wingless-related integration site"
WT	Wild-type
Zeb1	Zinc finger E-box-binding homeobox 1
Zeb2	Zinc finger E-box-binding homeobox 2
Zfhx1b	Zinc finger homeobox 1b

List of figures

List of tables

Summary in English

Cell fate decisions during embryogenesis and in proliferation and regeneration competent adult tissues and organs are orchestrated by large numbers of regulators including transcription factors, non-coding RNAs, epigenetic modifiers and longer-range genomic interactions, each downstream of extrinsic signals and their intracellular signal transduction effector proteins. Zeb2 is a DNA-binding transcription factor that steers in such context many cellular processes during early and late embryogenesis, and knowledge is emerging about its role in various adult tissues/organs as well, especially in challenge conditions like tissue organ failure. Mutations in one allele of *ZEB2* in humans cause Mowat-Wilson syndrome (MOWS) characterized by severe intellectual disability and, varying from patient to patient, epilepsy, Hirschsprung disease and other defects. In the mouse Zeb2 has been show, including by our laboratory, to regulate different aspects of central and peripheral nervous system (CNS, PNS, respectively) development as well as their post-natal functions. The same applies to embryonic and adult hematopoiesis, T-cell lineage development and function, and *e.g.* myelinogenesis in the CNS and (re)myelination by Schwann cells in the PNS. In addition, deregulation of the levels of Zeb2, *e.g.*. an abnormal increase, contributes (in this case) to many cancers and correlates with bad prognosis in humans, as supported by *Zeb2* overexpression mouse models.

The action mechanisms by which Zeb2, in concert with many of its emerging protein partners, steers neurogenesis and gliogenesis can be studied in cultured embryonic stem cells (ESCs) and the early CNS *in vivo*. In human ESCs, intact levels of ZEB2 were shown to be essential for effective neuroectoderm differentiation. The goals of this PhD project were to understand where, when and how Zeb2 functions (i) during mouse brain cortex development (using an *in vivo* approach) and (ii) in pluripotency and subsequent differentiation of mouse ESCs (*in vitro* approach).

Genetic inactivation of both alleles of *Zeb2* in early neural progenitors (using a Nestin-Cre approach), from dorsal telencephalon (Emx1-Cre) as well as in post-mitotic (NEX-Cre) upper layer neurons in the forming cortex in the embryonic forebrain causes a shift forward in the timing of first embryonic neurogenesis followed by embryonic and early post-natal gliogenesis. This leads to the expansion of the upper layers of the cortex at the expense of the (earlier born) cells of the deeper layers. Using gene expression profiling followed by validation, which is an essential part of this work and this PhD project, we identified neurotrophin-3 (Nt3) and fibroblast growth factor-9 (Fgf9) as extrinsic factors whose levels were significantly increased in the upper layer cells of the *Zeb2* knockout (KO) embryonic brain cortex as compared to control. In this way, Zeb2 regulates the generation of subsequent waves of neurogenesis (steered by Nt3 that acts on its Trk receptor complex in the progenitor cells and promotes neurogenesis) and of gliogenesis (via Fgf9 that acts on its receptors

there) emanating from the Zeb2-negative progenitor cells in the cortex. Importantly, this work identified Zeb2 as the first transcription factor acting in a cell non-autonomous fashion in cortex development.

In the second part of this PhD research the focus was on the role(s) and action mechanism(s) of Zeb2 in pluripotency and during differentiation of ESCs. For this, Zeb2-deficient (KO) ESCs were established. Using these novel, unique tools in combination with omics, this PhD project was able to add important new functional and mechanistic insight to previous findings. We discovered that Zeb2 is critical for exit from the epiblast state in mouse ESCs and links the pluripotency network and DNA-methylation with irreversible commitment to neural and general differentiation. In particular, we show that *Zeb2* KO ESCs display impaired differentiation in embryoid bodies by stalling in an epiblast-like state. Using RNA-Seq we further conclude that Zeb2 mainly acts here as a transcriptional repressor for many genes, either directly or indirectly, in differentiating conditions. Epithelial-mesenchymal transition (EMT), pluripotency, lineage commitment and DNA-(de)methylation genes are deregulated in *Zeb2* KO embryoid bodies. By using methylome analysis, we demonstrate these mutant cells cannot maintain their initially acquired DNA-methylation marks in neural-stimulating condition and do not effectively downregulate *Oct4*, *Nanog* and *Tet1* in differentiation conditions. Tet1 knockdown by RNA interference partially rescues the impaired differentiation of these KO cells. Another part of the PhD project investigated, starting again from the RNA-Seq data, the neuronal inhibitory *REST* gene, one of the genes whose expression was not silenced in *Zeb2* KO ESCs, unlike in normal ESCs. Like for the Tet1 studies, stable REST knockdown lines were established in the *Zeb2* KO background. The neural differentiation capacity of such cells line was also partially restored, indicating that *REST* is another important Zeb2-dependent gene in ESCs. In a last set of experiments, which provide the strong basis for future structure-function analysis of Zeb2 (and hence insight into its functional domains) in this ESC system, the neurodevelopmental potential of Zeb2 Smad-binding domain (SBD) mutant cells was documented. We observed that deletion of the SBD from Zeb2 had a positive effect on neural differentiation, indicating that this Zeb2 domain, and hence possibly its interaction with Smads, co-determines Zeb2's neural-inducing activity.

Samenvatting in het Nederlands

De beslissingen over het lot van cellen in het embryo en in proliferatie- en regeneratiecompetente adulte weefsels en organen worden georkestreerd door grote aantallen regulatorische molecules, inbegrepen transcriptiefactoren, niet-coderende RNAs, epigenetisch actieve eiwitten, en langere-afstand genomische interacties, elk stroomafwaarts van extrinsieke signalen en hun intracellulaire effectoreiwitten van hun signaaltransductie. Zeb2 is een DNA-bindende transcriptiefactor die in dergelijke context meerdere cellulaire processen stuurt gedurende de vroege en late ontwikkeling van het embryo. Onze kennis groeit over zijn rol, ook in verschillende adulte weefsels/organen, vooral recent ook in uitdagende omstandigheden zoals weefsel- of orgaanfalen. Mutaties in één *ZEB2* allel bij de mens veroorzaakt Mowat-Wilson syndroom (MOWS), gekarakteriseerd door ernstige verstandelijke beperkingen en, variërend van patiënt tot patiënt, epilepsie, ziekte van Hirschsprung en andere defecten. Er is aangetoond, ook door ons team, dat Zeb2 in de muis meerdere aspecten van centraal en perifeer zenuwstelsel (respectievelijk CZS, PZS) ontwikkeling, maar ook postnatale functies ervan, reguleert. Dit is ook het geval voor embryonale en adulte hematopoïese, T-cel ontwikkeling en functie en b.v. myelinogenese in het CZS en (re)myelinatie door Schwanncellen in het PZS. Bovendien is de deregulatie van de Zeb2 hoeveelheden, zoals een abnormale toename, een belangrijke factor (in dit voorbeeld) in meerdere kankers bij de mens en correleert het daar met barslechte prognose, wat bevestigd wordt door onderzoek met *Zeb2* overexpressie muismodellen.

De door Zeb2, samen met vele van zijn groeiende lijst partnerwitten, gehanteerde actiemechanismen die neurogenese en gliogenese sturen, kunnen worden bestudeerd in culturen van embryonale stamcellen (ESC) alsook in het vroege CZS *in vivo*. In humane ESC zijn normale niveaus van ZEB2 essentieel voor effectieve mesendoderm differentiatie. De doelstelling van dit doctoraatsproject was om beter te begrijpen waar, wanneer en hoe Zeb2 functioneert (i) in de ontwikkeling van de hersencortex van de muis (door *in vivo* studies) en (ii) in pluripotentie en daaropvolgende differentiatie van muis ESC (door een *in vitro* benadering).

Genetische inactivatie van beide allelen van *Zeb2* in vroege neurale voorlopercellen (door middel van een Nestin-Cre benadering), in de dorsale voorhersenen (Emx1-Cre) en ook in post-mitotische (NEX-Cre) neuronen van de opperste lagen van de cortex in de embryonale voorhersenen veroorzaakt een vroegere start van eerst embryonale neurogenese gevolgd door vroegere embryonale zowel als vroeg-postnatale gliogenese. Dit leidt tot een expansie van de bovenste lagen van de cortex ten koste van de (vroeger geboren) cellen van de diepere lagen. Studies van de genexpressie gevolgd door validatie, een essentieel deel van dit onderzoek en dit doctoraatsproject, identificeerden neurotrophin-3 (Nt3) en fibroblast groeifactor-9 (Fgf9) als extrinsieke factoren waarvan de dosage sterk was verhoogd in de bovenste lagen van de *Zeb2* knockout (KO) embryonale hersencortex, waar Zeb2

uitsluitend aanwezig is. Hierdoor reguleert Zeb2 de generatie van opeenvolgende golven van neurogenese (gestuurd door Nt3 dat zijn effect uitoefent via de Trk receptor complexen in de progenitorcellen, en neurogenese bevordert) en van gliogenese (via Fgf9 dat daar bindt op zijn receptors) die uitgaan van de Zeb2-negatieve progenitorcellen in de cortex. Zeer belangrijk is dat dit werk Zeb2 identificeerde als de eerste transcriptiefactor in hersencortex ontwikkeling die op een cel niet-autonome manier zijn effect uitoefent hierop.

In het tweede deel van dit doctoraatsonderzoek ging onze focus naar de rol(len) en actiemechanisme(n) van Zeb2 in pluripotentie en bij differentiatie van ESC. Hiervoor werden eerst Zeb2-deficiënte (knockout, KO) ESC aangelegd. Gebruik makend van deze unieke cellen en "omics" benaderingen kon dit doctoraatsonderzoek belangrijke nieuwe functionele en mechanistische inzichten toevoegen aan onze vroegere resultaten. Wij ontdekten dat Zeb2 kritisch is voor de uitstap van de epiblast status in muis ESC en dat Zeb2 het pluripotentie netwerk en de DNA-methylatie koppelt aan de onomkeerbare beslissing om neurale en algemene differentiatie te initiëren. Meer bepaald hebben we aangetoond dat *Zeb2* KO ESC defecten vertonen in hun differentiatie, getest in zgn. embryoid bodies, omdat zij vast blijven zitten in een epiblast-achtige status. RNA-Seq toonde verder aan dat Zeb2 vooral functioneert als transcriptionele repressor voor meerdere genen, hetzij direct of indirect, in ESC differentiatie. Voorbeelden van gedereguleerde genen in *Zeb2* KO embryoid bodies zijn die in epitheliaal-mesenchymale transitie (EMT), pluripotentie, differentiatie en DNA-(de)methylatie. Analyse van het methyloom toonde aan dat deze gemuteerde cellen hun initieel verworven DNA-methylatie signatuur niet kunnen behouden in neurale-stimulatie condities. Zij falen ook in de effectieve neerregulatie van *Oct4*, *Nanog* en *Tet1* in differentiatiecondities. Tet1 knockdown door RNA-interferentie redt partieel de defectieve differentiatie van deze KO cellen. Een ander deel van dit doctoraatsproject onderzocht ook, opnieuw startende van de RNA-Seq resultaten, het neuronaal-inhibitorische REST gen, één van de genen waarvan de expressie niet werd neergereguleerd in *Zeb2* KO ESC, in tegenstelling tot normale ESC. Zoals in de Tet1 studies werden stabiele knockdown cellijnen aangelegd in de *Zeb2* KO achtergrond. De neurale differentiatie van dergelijke cellen werd hierdoor ook partieel hersteld, hiermee aantonend dat *REST* een ander belangrijk Zeb2-afhankelijk gen is in ESC. In een laatste reeks van experimenten die de sterke basis leggen voor toekomstig structuur-functie onderzoek van Zeb2 (en hiermee inzicht in zijn functionele domeinen), werd het neuro-developmenteel potentieel van Zeb2 Smad-bindend domein (SBD) mutante cellen gedocumenteerd. Deletie van het SBD van Zeb2 had een positief effect op neurale differentiatie. Dit duidt op een belangrijke rol van dit Zeb2 domein, en mogelijk dus de interactie van Zeb2 met Smads, dat mee de neuraal-inducerende activiteit van Zeb2 bepaalt.

1 Chapter 1: General introduction

1.1 Early embryonic development and the first cell fate decisions

The first days of development after fertilization of the mouse oocyte involve cleavage of the embryo, i.e. a series of slow cell divisions, and subsequently eventually producing a ±32-cell, cavity-lacking embryo with cells at the outside (the trophectoderm) and the inside (inner cell mass, ICM). Immediately after, we can identify three major cell types that form the trophectoderm (TE), and within the ICM the primitive endoderm (PE, also named hypoblast) and epiblast (Epi, also named primitive ectoderm), respectively (for a schematic representation of the first stages of development, see Fig.1). The first molecular events in cell specification, but with retention of pluripotency of each blastomere, have been proposed to occur already in the 4-cell stage embryo, wherein heterogeneously distributed Prdm14 (PR Domain Containing 14), a member of the PRDI-BF1 and RIZ homology domain containing (PRDM) family of transcriptional regulators, together with Carm1 (Co-activator-Associated Arginine Methyltransferase 1), selectively directs the cells to become ICM by inducing H3 arginine methylation, coinciding with upregulation of Nanog and Sox2 mRNA in these cells (Burton *et al.*, 2013; Torres-Padilla *et al.*, 2007). It are then the next two rounds of cell division (from 8 to 16 and from 16 to32 cells) that produce two distinct cell types that will become TE and ICM, respectively. The future trophoblast cells occupy ultimately the outer layer of this embryo (see above), are larger than the rest of the embryonic cells and express high levels of Cdx2 mRNA, whereas future ICM cells are smaller and group together in the center of the embryo and loose Cdx2 mRNA (Jedrusik *et al.*, 2008). TE specification thus occurs during transition from the 8 to 16-cell stage; it is controlled by the transcription factor Tead4 (TEA Domain Family Member 4) whose activities require two co-activators, Yap and Taz, regulated by the core Hippo signaling pathway kinases Lats1/2. The active Hippo pathway inhibits Yap and Taz activity and prevents them from reaching the nucleus where Tead4 is located, which on its turn prohibits *Tead4* and *Cdx2* activation (Nishioka *et al.*, 2009).

In the first step towards Epi versus PE lineage segregation a subset of (future PE) cells acquires high FgfR2 levels while (the future) Epi cells induce *Fgf4* (Morris *et al.*, 2013). Prior to definitive lineage allocation, around E3.5 (±64 cells), the “salt and pepper” presence of Gata6 (in PE cells) and Nanog (in the Epi cells, at variable levels) becomes visible (Rossant *et al.*, 2003). This differential gene expression is reinforced by Fgf signaling (Kang *et al.*, 2013). Segregation of the Epi from the PE cells

occurs between E3.5 and 4.5. The late blastocyst (E4.5, 100-120 cells) is, as mentioned briefly above, composed of trophoblast (Cdx2, Eomes and Gata3 positive (+)) that surrounds the epiblast (Nanog+ and Oct4+), with the latter being separated from the blastocoel cavity by a thin layer of PE cells (Gata6+, Sox7+ and Sox17+).

The mouse embryo hatches and transforms into an egg cylinder whilst implanting in the uterus shortly after E4.5. The TE serves and mediates implantation and gives rise to the extraembryonic ectoderm and the ectoplacental cone (which will contribute to the placenta). The PE gives rise to the parietal endoderm and the visceral endoderm, which will later give rise to the endoderm of the visceral yolk sac. The early epiblast undergoes a global reshaping to ultimately form a cup-shaped polarized epithelium that will give rise to the somatic tissues and germ cell lineage of the embryo proper (reviewed in Arnold and Robertson, 2009).

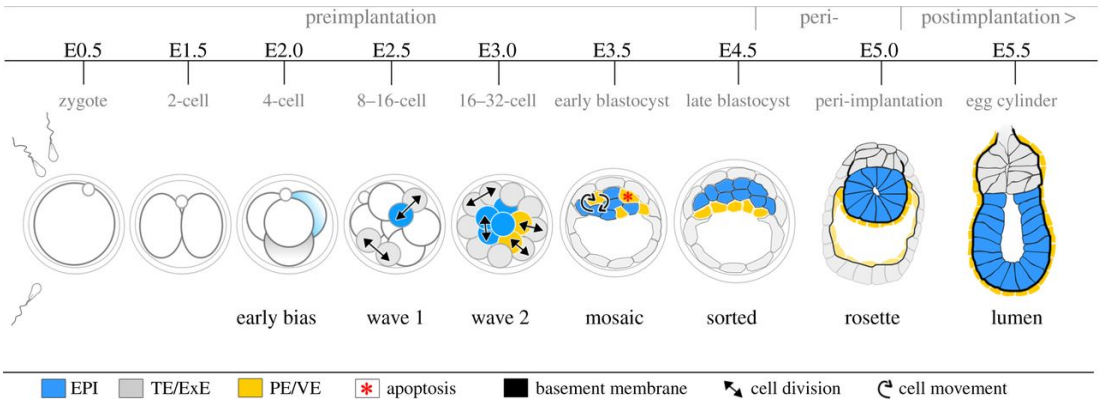

Figure 1. Overview of early mouse development.

Embryonic and extraembryonic cells are specified in the preimplantation embryo by two cell fate decisions. In the first cell fate decision, waves of cell divisions create inside and outside cells. Outside cells give rise to extraembryonic trophectoderm (TE), while inside cells form the pluripotent inner cell mass (ICM). In the second cell fate decision, cells of the ICM are segregated into the extraembryonic PE and the pluripotent epiblast (EPI) that will later give rise to all tissues of the body. These fate decisions are influenced but not determined by heterogeneity between individual cells within the embryo that is established by the 4-cell stage (shown by different shading of cells). At E4.5, the embryo initiates implantation, cell proliferation rapidly enhances and over the next 24 h the implanted embryo invades the maternal tissues, continues to display cell proliferation and transforms into an egg cylinder. This new form serves as a foundation for EPI patterning, laying down the body axis and establishment of the germ layers. The parietal endoderm and the accompanying and the trophoblast giant cells originating from the mural TE, as well as the ectoplacental cone derived from the polar trophectoderm, are omitted from the drawing of the E5.5 embryo. (Firgue and legend taken from: Bedzhov et al., 2014)
Abbreviations: ExE, extraembryonic ectoderm; PE, primitive endoderm; VE, visceral endoderm.

Starting from E6.5 the embryo undergoes gastrulation in which the three ultimate germ layers are generated and re-positioned by massive cell migration, laying down the primitive body plan of the embryo. Ectoderm, mesoderm and definitive endoderm (emanating from the mesendoderm) then give rise to all embryonic cells from which the embryo and further extra-embryonic tissues will develop.

1.2 Embryonic stem cells

In this section we address the current operational definition(s) of pluripotency, *in vivo* as well as *in vitro*, and we discuss various factors that influence the pluripotent state(s) of stem cells of the early embryo.

1.2.1 ESC discovery

The establishment of the first embryonic stem cell (ESC) lines was preceded by many years of study of embryonal carcinoma (EC) cells, the undifferentiated stem cells of teratocarcinomas, and their *in vitro* and *in vivo* differentiation capacities. Such studies, which actually led to the formulation of the *stem cell theory of cancer*, also helped to establish the initial and subsequently successful, ideal cell culture conditions for ESCs. Teratocarcinomas are malignant tumors, which indeed contain ECs and a differentiated component that can include cells representative of one of the three embryonic germ layers. Teratomas are tumors composed of somatic tissues only, hence are devoid of ECs and therefore these tumors are benign. The first experimental proof that ECs are multi-potent with regard to their developmental potential was obtained by Kleinsmith and Pierce back in the early 60s. Limiting dilution and ultimately single-cell transplant experiments demonstrated that ECs were able to reconstitute teratocarcinomas containing various differentiated types of cell. The ECs had the capacity to self-renew in culture, were pluripotent (*i.e.* could produce all somatic lineages upon injection in an immune-compatible or immune-deficient host) and also tumorigenic and cytogenetically abnormal (Kleinsmith and Pierce, 1964). A few years later the first stable EC cell lines were established and their multi-lineage differentiation potential was further demonstrated *in vitro* (Kahan and Ephrussi, 1970). The teams of Solter and Stevens showed that early embryos transplanted to mouse testis could generate teratomas. This collectively proved that an early embryo contains a pluripotent cell population and that - under the right cell culture conditions - these cells could be expanded *in vitro* (Solter *et al.*, 1970; Stevens, 1970).

In 1981 two groups reported the derivation of mouse ESCs from pre-implantation embryos using the same culture conditions as for the EC lines, and using mouse embryonic fibroblasts (MEFs)

as a feeder layer in a cell culture medium, which was supplemented with serum (Evans and Kaufman, 1981; Martin, 1981). The authors showed that these ESCs could be propagated *in vitro*, were karyotypically normal (unlike most EC lines) and could contribute to various tissues in ESC: acceptor chimeric embryos, and also formed teratomas when injected into mice.

1.2.2 States of pluripotency

Pluripotent embryonic cells have the capacity to differentiate into any of the three germ layers of the embryo proper. *In vivo*, in the mouse, cells of the epiblast, in a time window between E3.5 till E7.5, display pluripotent potential (reviewed in Boiani and Schöler, 2005; Nichols and Smith, 2012). ICM cells of the E4.0-E4.5 embryo represent the often-referred-to **ground-state** of pluripotency with the least restricted lineage potential characterized by high, uniform expression of pluripotency genes (at least when assessed by population analysis, *i.e.* not at the single-cell level) and - in female embryos - reactivation of the paternal X chromosome (Mak *et al.*, 2004; Nichols *et al.*, 2009; Silva and Smith, 2008). These ground-state pluripotent cells *in vivo* are naïve, *i.e.* they have unbiased developmental potential. Upon implantation, shortly after E4.5, lineage-specific genes start to be expressed and a global increase in DNA-methylation is observed, which indicates lineage priming hence referred to as the primed pluripotency state (Borgel *et al.*, 2010). Pluripotency is eventually silenced around E7.5 (Osorno *et al.*, 2012). Interestingly, although the molecular signatures of the epiblast cells change during developmental progression between E3.5 and E7.5, all these cells retain their 3-lineage differentiation potential, in fact showing that a broader spectrum of pluripotent state may exist *in vivo*, which embryologists increasingly are trying to catch experimentally and characterize by *a.o.* extensive transcriptomics, including *in vitro*.

Mouse ESCs (Evans and Kaufman, 1981; Martin, 1981), and later the isolated epiblast stem cells (EpiSCs) (Brons *et al.*, 2007; Tesar *et al.*, 2007) and embryonic germ cells (EGCs) (Leitch *et al.*, 2010; Matsui *et al.*, 1992) are the three embryonic pluripotent cell types that currently can be maintained *in vitro* (see Fig. 2). ESCs and EGCs have similar transcriptional and epigenetic profiles and rely on the same signals and intracellular signaling pathways for maintaining their capacity of self-renewal as undifferentiated cells (Leitch *et al.*, 2013), whereas EpiSCs are more primed and require distinct culture conditions to maintain pluripotent potential (Brons *et al.*, 2007; Sugimoto *et al.*, 2015; Tesar *et al.*, 2007). Interestingly, by simply manipulating cell culture conditions it is possible to derive EpiSCs from ESCs (Tosolini and Jouneau, 2015). It has also been shown that conversion of post-implantation epiblast stem cells to ESCs is possible in a period of 14-35 days in the presence of serum and LIF (Bao *et al.*, 2009; Gillich *et al.*, 2013). However in order to obtain ESCs from EpiSCs with high

efficiency, single (Klf4, Guo *et al.*, 2009) or - for more rapid conversion - double (Klf2 with Prdm14, Gillich *et al.*, 2012) transgene overexpression was needed.

ESCs cultured on feeder cells in growth medium containing serum (in this thesis text referred to as “serum ESCs”) – at the population level – are pluripotent. Detailed examination showed however high heterogeneity in terms of gene expression, and oscillation between naïve and primed states of pluripotency within such ESC populations (Abranches *et al.*, 2013), believed to be due to simultaneous activation of multiple signaling pathways that together destabilize the naïve state and make the cells more prone to differentiation (reviewed in Nichols and Smith, 2009).

Recently it was shown that ESCs cultured in chemically defined medium (referred to as “2i” in this manuscript, see for more details below), using the respective inhibitors of MEK and GSK3, have their pro-differentiation pathways inhibited, stabilize their naïve state and strongly resemble ground-state cells of the E4.0-E4.5 epiblast. These “2i” cells are therefore named “ground-state” ESCs (Ying *et al.*, 2008). It was shown that total numbers of differentially expressed genes are the lowest between 2i ESCs and E4.5 embryonic epiblast, whereas the period between E3.75 and E4.5 allows the most efficient ESC line generation, and numbers of colonies obtained are proportional to the epiblast size. In contrast, single cells from early blastocysts E3.25-E3.5 rarely give rise to ESC lines (Boroviak *et al.*, 2014). Pluripotency is a functional or operational definition that likely covers more than one stage of early embryonic development that can be caught *in vitro*, and certainly applies to ground-state and primed ESCs, and EpiSCs. The state of pluripotency should always be specified, especially in the context of studies of the gene regulatory networks in *e.g.*. ESCs and early cell fate decisions in the vertebrate embryo.

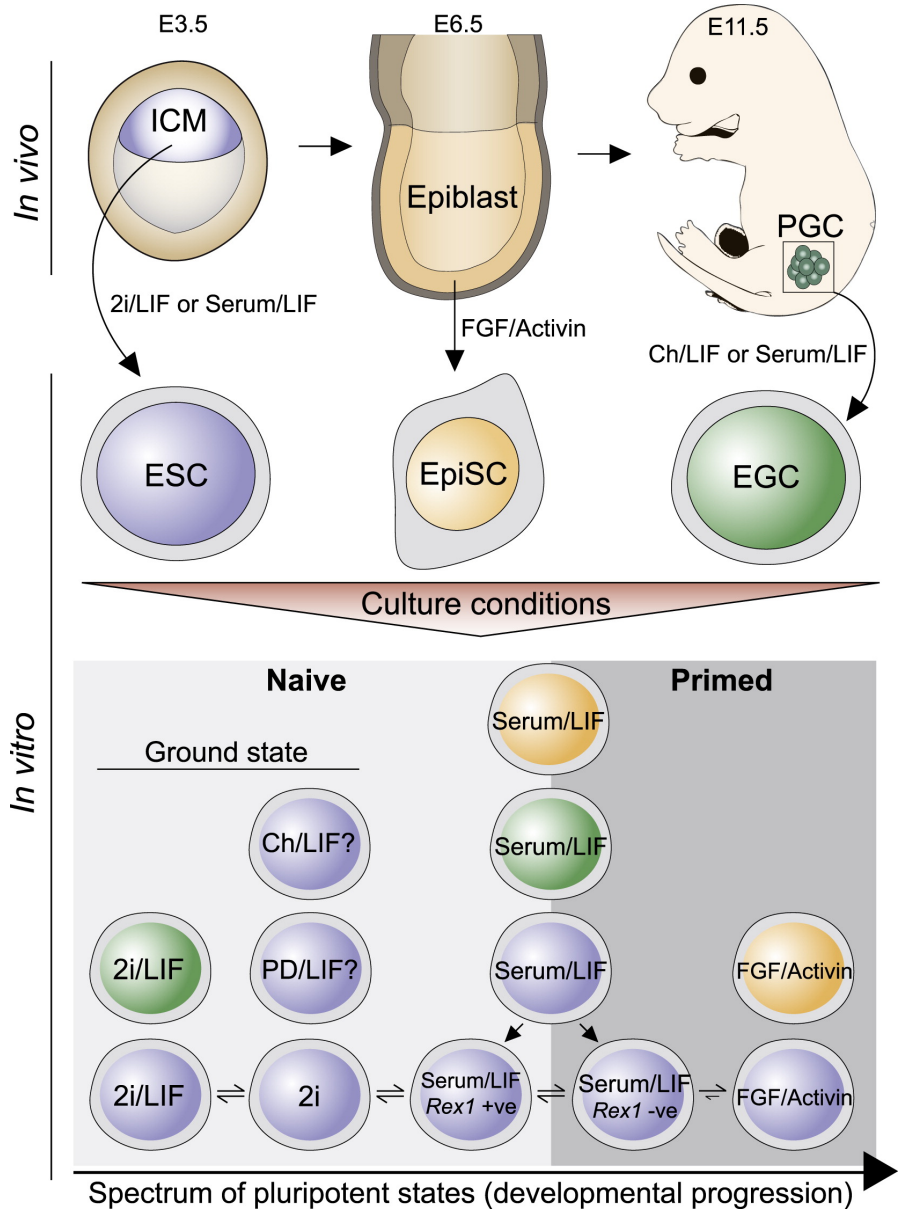

Figure 2. Embryonic origin and spectrum of pluripotent stem cell states.

The pluripotent cells of a blastocyst between E3.5 and E4.5 can give rise to functionally naïve ESCs (blue). Between E5.5 and E8.0 post-implantation epiblast can establish EpiSCs (orange), which occupy a primed pluripotent state. Additionally, primordial germ cells (PGCs), which are the founders of the germline lineage, can give rise to naïve EGCs (green), which are highly comparable to ESCs. Depending on the culture/derivation conditions these pluripotent stem cells occupy discrete molecular states that can be broadly classed as naïve or primed. The most optimized state of naïve pluripotency, which closely recapitulates the naïve epiblast cells of the blastocyst, is termed ground state. An interchangeable spectrum of pluripotent states may arise that ranges from ground state to primed pluripotency. The state of pluripotency adopted *in vitro* is primarily dictated by the combination of extrinsic signals in the culture environment rather than the developmental source of the pluripotent cells.
Other abbreviations: CH, Chiron, CHIR99021*; PD,* PD0325901(Figure and legend taken from: Hackett and Surani, 2014)

1.2.3 Extrinsic signaling pathways in regulation of ESC states

For schematic representation of the extrinsic signaling pathways in ESCs, see Fig.3.

Ground-state pluripotency is displayed in the early embryo for a short period of time, for it is rapidly destabilized by intrinsic and extrinsic signals. Cells of the epiblast acquire new molecular signatures with every cell division with which they seem to progressively lose their potency.

As mentioned above, ESCs can be maintained as self-renewing undifferentiated colonies for an unlimited period of time. Leukemia inhibitory factor (LIF) is the key extrinsic factor that sustains ESC self-renewal and propagation as undifferentiated cells (Smith *et al.*, 1988; Williams *et al.*, 1988). LIF binds to a high-affinity receptor complex composed of a low-affinity LIF binding chain (LIFR, also known as CD118) and a high-affinity converter subunit (Gp130) leading to phosphorylation and activation of the downstream, cytoplasmic transcription factor STAT3 (Yoshida *et al.*, 1994). Phosphorylated STAT3 accumulates in the nucleus and subsequently regulates expression of a large number of pluripotency/self-renewal genes including *Klf4, Tfcp2l1, Gbx2* and *c-Myc* (Cartwright *et al.*, 2005; Hall *et al.*, 2009; Martello *et al.*, 2013; Tai and Ying, 2013). Classic ESC culture conditions require LIF (produced by the feeder cells and/or supplemented as recombinant, bio-active cytokine in the medium) and serum. BMP4 that acts via Smad proteins to induce expression of downstream *Id* (Inhibitor of differentiation) family genes is a crucial component within serum added to these cultures. In fact, feeder-cell free ESC cultures supplemented only with both LIF and BMP4 retain their pluripotent and self-renewing capacity even in the absence of feeder cells. Upon BMP4 withdrawal (hence, in presence of LIF only) ESCs can acquire a neural phenotype, whereas removal of LIF (only leaving BMP4 present in the medium) results in non-neural differentiation (Finley *et al.*, 1999; Ying *et al.*, 2003).

Pluripotency of ESCs seems to depend on the balance between pro-self-renewal and pro-differentiation signals. ESCs themselves produce extrinsic factors that destabilize their ground state. One such factor is Fibroblast growth factor 4 (Fgf4). Stimulation by FGF leads to activation of the downstream Erk (Extracellular-signal-regulated kinase, also known as MAPK) pathway, which is believed to be an important trigger causing ESCs to switch from self-renewal to lineage commitment. Knockout of either *Fgf4* or *Erk1/2*, or chemical inhibition of the Fgf-Mek-Erk pathway, helps to maintain ESCs in their uncommitted state (Kunath *et al.*, 2007; reviewed in Lanner and Rossant, 2010). Activation of WNT signaling leads to inhibition of Gsk3 and – as a consequence – to increased stability and nuclear accumulation of the downstream effector of canonical Wnt signaling, β-catenin. This β-catenin enhances Oct4 activity (Kelly *et al.*, 2011) and – via the DNA-binding transcription factor Tcf3 – regulates expression of *Esrrb*, one of the key pluripotency genes (Martello *et al.*, 2012). Wnt is essential for ESC self-renewal as it prevents the cells from transiting to the EpiSC state (ten Berge *et al.*, 2011).

Conversely, Wnt inhibition allows efficient derivation of EpiSCs (Sugimoto *et al.*, 2015). Also, chemical inhibition of GSK3 activity mimics the effects of WNT signaling activation and supports self-renewal (Ogawa *et al.*, 2006; Sato *et al.*, 2004). Chemical inhibition of both FGF (via MEK inhibitor PD0325901) and WNT (via GSK3 inhibitor CHIRON, CHIR 99021) pathways, together with added LIF, is the basis for the aforementioned "ground-state" ESC medium known as "2i" (Ying *et al.*, 2008).

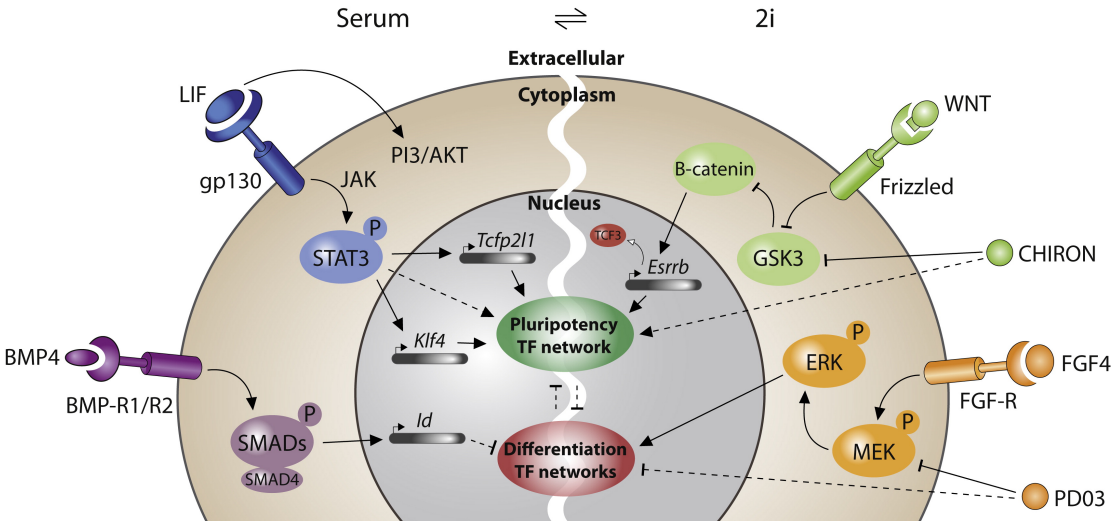

Figure 3. Extrinsic signaling pathways that feed into reinforcing or antagonizing naive pluripotency.

Simplified representation of various signaling cascades that affect self-renewal. Filled arrows indicate activation, whereas bars show inhibition or blockade of target activity. A solid line implies a direct or known downstream target and a dashed line indicates an indirect or inferred effect. Clockwise: BMP4 is present in serum and functions via SMADs to activate *Id* genes. LIF signaling affects many pathways but primarily acts via JAK-mediated phosphorylation of STAT3, which activates *Tcfp2l1* and *Klf4*. Canonical WNT signaling blocks GSK3 activity leading to stabilization of β-catenin, which in turn abrogates TCF3-mediated repression of pluripotency genes including *Esrrb*. CHIRON closely mimics WNT signaling by inhibiting GSK3. FGF signaling activates the MAPK pathway leading to phosphorylation of MEK, which in turn phosphorylates and activates ERK. Activated ERK promotes transition to a "primed" state, which is therefore blocked by the MEK inhibitor PD03.(PD0325901) (Figure and legend taken from: Hackett and Surani, 2014)

1.2.4 Transcription factors relevant for acquisition and maintenance of pluripotency and their networks

For schematic representation of known key transcription factors (TFs) important for self-renewal of serum and 2i ESCs, see Fig. 4.

TFs recognize specific DNA sequences present in the proximal gene promoter regions as well as in distal regions located often (up to) hundreds of kilobases away from the transcription start

sites. They recruit (or release) the transcriptional machinery of the cell and/or bind chromatin-modifying proteins to allow access of other TFs and TF complexes (Fuda *et al.*, 2009; Li *et al.*, 2007a).

Oct4, Sox2 and Nanog are considered the core TFs in the pluripotency network of ESCs. They are induced during the establishment of pluripotency *in vivo* and together guard the pluripotent state *in vitro* (Chambers and Smith, 2004; Chambers *et al.*, 2003; Nichols *et al.*, 1998). Both knockdown (by more than 50%) and overexpression (by more than 150%) of Oct4 lead to ESC differentiation (Masui *et al.*, 2007; Niwa *et al.*, 2000). Surprisingly, lowering Oct4 levels by half (using Oct4+/- ESCs), seems to stabilize naïve pluripotency (Karwacki-Neisius *et al.*, 2013). This shows that Oct4 levels need to be precisely controlled in ESCs, including for maintaining robust self-renewal of these cells. *Sox2 null* embryos die shortly after implantation (Avilion *et al.*, 2003). Sox2 very often acts in concert with Oct4 to regulate target gene expression in stem cells. Moreover, Oct4 and Sox2 co-regulate and enhance each other's expression in ESCs (Chew *et al.*, 2005; Masui *et al.*, 2007). Without Nanog, pluripotency does not develop *in vivo* (Silva *et al.*, 2009). Surprisingly, Nanog is dispensable for the *in vitro* pluripotency maintenance and knockout of *Nanog* only slightly increases the ESC propensity to differentiate (Chambers *et al.*, 2007). Although it is not crucial for pluripotency maintenance, Nanog overproduction bypasses the need for LIF to self-renew, which emphasizes its role in the pluripotency network (Chambers *et al.*, 2003; Mitsui *et al.*, 2003). Oct4 and Sox2 are in the center of the ESC transcriptional network where, together with Nanog, they form a positive auto-regulatory loop that enhances/stabilizes their mRNA expression. They were also shown to co-operate with a large number of claimed "non-essential", but still pluripotency-related genes that form the *extended* network of ESC pluripotency and reinforce the ESC ground state. Below, a number of such genes are briefly discussed.

Esrrb is a direct target of Nanog and can substitute for Nanog in ESCs. It was also shown to act downstream of Gsk3/β-catenin and β-catenin/Tcf3. Forced production of Esrrb resembles both Gsk3 inhibition and *Tcf3* knockout phenotypes, and such Esrrb bypasses the need for cytokines to maintain the undifferentiated state (Festuccia *et al.*, 2015; Martello *et al.*, 2012). ***Tfcp2l1*** is a target of both LIF/Stat3 and 2i-mediated self-renewal. Forced production of Tfcp2l1 can recapitulate the self-renewal promoting effect of LIF or either of the 2i components. Hence, Tfcp2l1 has been proposed to play a critical role in maintaining ESC identity (Martello *et al.*, 2013; Ye *et al.*, 2013). ***Klf4*** also plays a role in preventing ESC differentiation by direct regulation of *Nanog*. Knockdown of Klf4 induces ESC differentiation, whereas overproduction of Klf4 reinforces ESC self-renewal and abrogates the requirement for LIF (Li *et al.*, 2004; Zhang *et al.*, 2010). Another member of the Klf family, ***Klf2*** – transcriptionally induced by Oct4 – has been shown to promote self-renewal (Hall *et al.*, 2009). At the transcriptional level, **Sall4** regulates *Oct4*. Knockdown of Sall4 in ESCs resembles Oct4 knockdown and results in a shift from self-renewal to trophoblast differentiation (Zhang *et al.*, 2006). The aforementioned Prdm14 plays a dual function is ESC identity maintenance by antagonizing Fgf-induced

cells differentiation on the one hand and inhibiting *de novo* methylation mediated by DNA-methyltransferases on the other hand (reviewed in Nakaki and Saitou, 2014; Yamaji *et al.*, 2013).

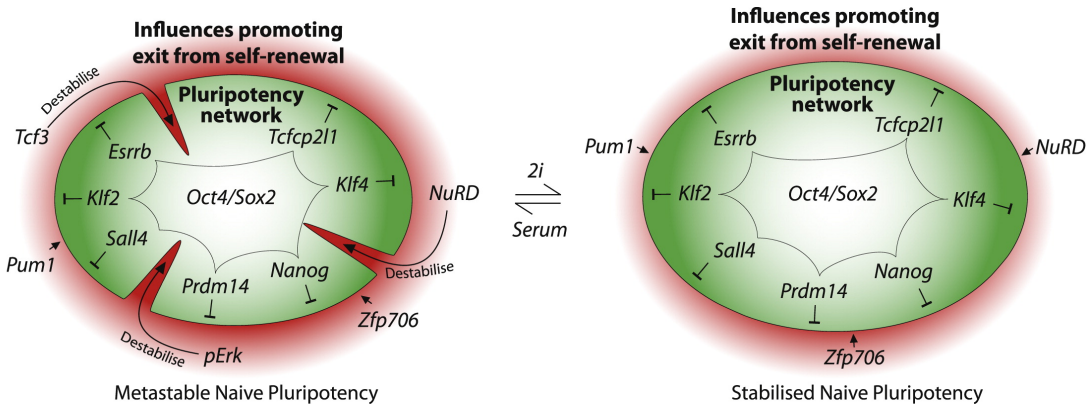

Figure 4. Genetic networks for maintenance of naive pluripotency.

The pluripotency network includes essential core components (*Oct4/Sox2*) and multiple ancillary factors (shown in green) that collectively form a self-organizing circuitry. Ancillary factors generally act to buffer the network against fluctuations and perturbations that promote exit from self-renewal (shown in red). In serum (left), ancillary factors are heterogeneous among the population and thus render individual ESCs in varying states of susceptibility to inherent influences that promote exit from self-renewal such as the FGF/ERK pathway or NuRD complex activity (shown as canals in the ancillary bubble). When the pluripotency network is stabilized through blocking key differentiation influences and/or by directly reinforcing expression of ancillary factors (*e.g.* 2i/LIF, right), the influences that drive exit from self-renewal are mitigated and robust naïve pluripotency emerges. (Figure and legend taken from: Hackett and Surani, 2014)

To identify the components and fully understand the transcription network properties that control cell state, one should examine undifferentiated versus differentiated (here also naïve versus primed) cell states, integrate time-course multi-lineage differentiation data, perform large-scale genetic perturbation screens, investigate promoter/enhancer occupancy and connected histone signatures and/or perform computational modeling (which should be subsequently experimentally validated). In recent years a number of genome-wide studies addressed the complex global architecture of the transcriptional network of ESC populations. More recently, such studies are being re-addressed at the single-cell level. One of the first large-scale studies that addressed ESC self-renewal were those by Ivanova *et al.* (2006). They combined gene perturbation (using shRNA) with dynamic, global analyses of gene expression and showed that Esrrb, Tbx3 and Tcl1, as well as the previously identified Nanog, Oct4 and Sox2, are required for efficient self-renewal of ESCs. Furthermore, they demonstrated that at least two independent signaling pathways supporting self-renewal exist *in vitro*: one was found to respond to perturbation of Nanog, Oct4 and Sox2 but not of Esrrb, Tbx3, Tcl1 or Dppa4; the other responded to perturbation of Esrrb, Tbx3, Tcl1 or Dppa4, but not of Nanog, Oct4 and Sox2 (on top of a

large subset of genes responding to knockdown of all TFs used in the study). This work, as one of the first, revealed the complexity of cell fate regulation in ESCs and showed that pluripotency factors are globally involved both in gene activation and repression (Ivanova *et al.*, 2006).

In another study, Boyer *et al.* (2005) interrogated the core transcriptional network of Oct4, Sox2 and Nanog in human ESCs and for this they combined ChIP analysis with microarray-based mRNA expression profiling. The authors demonstrated that Oct4, Sox2 and Nanog co-occupy the promoters of a large number of genes that encode developmentally important homeodomain transcription factors, and that these regulators contribute to specialized regulatory circuits in ESCs. Importantly, the authors showed that the expression of *Oct4*, *Sox2* and *Nanog*, as proteins bound together to the promoters of their own genes, are interconnected via autoregulatory loops (Boyer *et al.*, 2005). Similar results were obtained by Loh *et al.* who used ChIP-PET in combination with knockdown and overproduction of *a.o.* Nanog and Oct4 in mouse ESCs (Loh *et al.*, 2006).

Chen *et al.* (2008) expanded the ESC network studies to 13 factors that were previously shown to play a role in self-renewal, in reprogramming and/or were a component of the LIF or BMP signaling pathways. They mapped the binding sites for Nanog, Oct4, STAT3, Smad1, Sox2, Zfx, c-Myc, n-Myc, Klf4, Esrrb, Tcfcp2l1, E2f1 and CTCF, and of two other transcriptional regulators, p300 and Suz12. They found that specific genomic regions extensively targeted by multiple TFs (called multiple transcription-factor-binding loci) function as ESC enhanceosomes. The authors showed that - among these 13 TFs - Nanog, Sox2, Oct4, Smad1 and STAT3 tend to co-occupy the same regions, as do members of a second, distinct group comprised of n-Myc, c-Myc, E2f1 and Zfx. Based on the associations between binding of the studied TF and mRNA expression level of the presumptive, closely located (and – in general – ideally TF-dependent) target genes (*i.e.* the expression data taken from Ivanova *et al.*, 2006; Zhou *et al.*, 2007), the authors constructed a transcriptional regulatory network model that integrates LIF and BMP signaling pathways with intrinsic factors in ESCs (Chen *et al.*, 2008).

Recently, new computational modelling defining the essential program for naïve pluripotency was applied and validated. This allowed to define the simplest version of the self-renewal network in ESCs: it comprises 16 interactions only, achieved by 12 components and three inputs, and can predict responses to compound and genetic perturbations with high accuracy. The essential network to propagate the naïve transcriptional state contains Oct4, Sox2, Nanog, Klf4, Klf2, Sall4, Esrrb, Gbx2, Tfcp2l1 together with inputs from Stat3, Tcf3 and Erk/Mek signaling pathways (Dunn *et al.*, 2014).

In conclusion, the transcriptional regulatory network of ESCs is very complex as it contains multiple interconnected autoregulatory loops together with nodes that allow smooth and reversible transitions between *e.g.* the ground and primed pluripotent state and ensures extraordinary self-renewal

robustness. Recent pioneering work in mouse ESCs at single-cell level, for example by the team of Timm Schroeder and colleagues (Basel), is re-addressing these. This work will also be discussed in the General Discussion section of this PhD thesis.

1.2.5 Non-coding RNAs important for stemness maintenance

Regulatory non-coding RNAs (ncRNAs) can be divided into two broad classes, small (<200 nucleotides (nt) long, including micro-RNAs (miRNAs or miRs), endogenous small-interfering RNAs (endo-siRNAs), Piwi-interacting RNAs (piRNAs)) and large/long non-coding RNAs (lncRNAs, >200 nt long). Biogenesis of the main subclasses of the small ncRNA depends on different proteins and protein complexes. miRNAs are generated via sequential post-transcriptional processing by the Drosha-DGCR8 and the RNaseIII;Dicer complexes, followed by assembly with Ago proteins into an RNA-induced silencing complex (RISC). The miRNA then guides the RISC complex to the target mRNA(s).

In contrast to miRNAs, the silencing effects of endo-siRNAs are dependent only on Dicer, and do not require Drosha-DGCR8 action. Finally, piRNAs require neither Drosha-DGCR8 nor Dicer, bind to PIWI-subfamily (instead of Ago-subfamily) Argonaute proteins, and silence their targets by mediating mRNA degradation, or possibly DNA-methylation (reviewed in: Huo and Zambidis, 2013). ncRNA biogenesis is reviewed in detail elsewhere (Kim *et al.*, 2009). Below, I present a selection of studies that highlight the importance of **ncRNAs** in stem cell biology.

It was shown that Dgcr8, which is required for miR processing, is essential for silencing of ESC self-renewal. In absence of Dgcr8, a global loss of miRs is observed and such ESCs do not silence the pluripotent genes efficiently and significantly delay the onset of lineage differentiation program (Wang *et al.*, 2007). Similarly, gene inactivation of *Dicer*, which is another important gene for miRNA and piRNA processing, leads to defects in small ncRNA biogenesis and ESC differentiation (Kanellopoulou *et al.*, 2005). Members of the miR-290/295 (in mouse ESCs) and miR-302-367 (in human ESCs) gene clusters are among the most abundantly expressed miRNA-encoding genes in ESCs. Overexpression of these miRNAs can partially rescue the phenotype caused by removal of Dgcr8, demonstrating their functional significance in the regulation of self-renewal of ESCs (Wang *et al.*, 2008). Interestingly, the core pluripotency factors Oct4, Sox2 and Nanog, and Tcf3, occupy a large fraction of miRNA promoters in ESCs (*i.e.* 20% of the annotated mammalian miRs). miRs whose promoters are bound by the core TFs are more abundant in ESCs as compared to differentiated cell types (for example MEFs) or neural precursor cells. Moreover, miRs inactive in ESCs (but active in other cell types) whose promoters are occupied by the core factors are often co-bound by the PRC2 complex associated with transcriptional

silencing, which suggests another level of complexity in gene expression regulation (Marson *et al.*, 2008).

lncRNAs do not encode functional proteins, however they are similar to mRNAs, for they are transcribed by RNA polymerase II and contain a 3' poly-A tail. lncRNAs can be encoded by sequences within introns of protein-coding genes (in sense and antisense direction), also overlap with protein-coding genes (in antisense direction) or be located in the intergenic regions of the genome (such RNAs are named large-intergenic RNAs, lincRNAs). The most comprehensive study to date addressing the role of lncRNAs in ESC self-renewal is the one by Guttman *et al.* and focuses on lincRNAs. They first identified over a thousand previously unknown conserved lincRNAs (Guttman *et al.*, 2009, 2010). Next, they performed a large-scale and systematic loss-of-function study, successfully targeting 147 (out of a total of 226 tested) lincRNAs that were active in ESCs, using a *Nanog* promoter-based luciferase reporter line as readout. The authors identified 15 lincRNAs whose knockdown affected ESC self-renewal. To investigate whether any of the 147 lincRNAs affected ESC lineage differentiation programs, the authors combined their loss-of-function data with previously published data sets of gene expression profiles for primitive endoderm, trophoblast and primitive ectoderm/neural ectoderm, respectively (Aiba *et al.*, 2009). Doing so, they identified 30 lincRNAs whose knockdown resulted in transcriptional changes similar to early lineage differentiation programs. This indicates that normally these lincRNAs have inhibitory effects on lineage commitment gene expression. Furthermore, the authors showed that 75% of the ESC-lincRNAs were bound by at least one of the pluripotency regulators, with a median of 3 TFs bound by each promoter, and at least 50% of these interactions were functional as knockdown of pluripotency TFs resulted in changes of the target lincRNA expression (ChIP data taken from Chen *et al.*, 2008). In addition, the authors demonstrated that many of the ESC-associated lincRNAs were bound by chromatin-modifying complexes such as PRC2, PRC1 and Jarid1b (Guttman *et al.*, 2011).

In conclusion, several lines of evidence support the concept that non-coding RNAs are "wired" in the self-renewal network of ESCs and co-control lineage commitment.

1.2.6 Divergent epigenetic landscapes of ground, primed and differentiating ESCs. Enzymes and protein complexes that modify the epigenetic changes.

DNA-methylation and hydroxylation, modifications of histones (especially methylation and acetylation) and chromatin remodeling complexes together shape the epigenetic landscape and add complex, interconnected and as such extra regulatory mechanisms to stem cell state control. Pluripotent cells have a specific epigenetic signature that changes when the cells enter differentiation. Below, I discuss

each (of these three) epigenetic modifications separately and give examples of how these modifications influence/reflect ESC states. I also summarize the epigenetic signatures of different pluripotent cell types.

For the summary of selected knockout and knockdown phenotypes in mouse and ESCs, see Table1 (p.37-38).

1.2.6.1 DNA-methylation

De novo DNA-methylation is controlled by DNA (cytosine-5)-methyltransferases (Dnmts) that together form a small family (Dnmt3a, Dnmt3b, Dnmt3l), members of which are highly active in ESCs and early embryos where they establish new methylation patterns. Another Dnmt, known as the maintenance methyltransferase Dnmt1, copies the methylation pattern onto daughter cells during somatic cell division (Okano *et al.*, 1999; Goll and Bestor, 2005). Dnmt3a, Dnmt3b together with Dnmt3l modify the epigenome of ESCs into a primed EpiSC-like state. Their genes are repressed by Prdm14 (Yamaji *et al.*, 2013). Conversely, upon ESC differentiation, *Dnmt3a*/*3b* expression increases and both encoded enzymes synergistically methylate the *Nanog* and *Oct4* promoter (Li *et al.*, 2007b). Gene inactivation of *Dnmt1* or *Dnmt3b* in the mouse results in perinatal death due to multiple developmental defects, whereas Dnmt3a-deficient animals survive till 1 month after birth (Li *et al.*, 1992; Okano *et al.*, 1999). Surprisingly, deletion of both alleles of all three *Dnmts* (*i.e. Dnmt1*, *3a* and *3b*) in mouse ESCs - causing loss of CpG methylation - does not affect their self-renewal or global chromatin architecture (Tsumura *et al.*, 2006). In contrast to mouse ESCs, simultaneous deletion of *Dnmt3a*, *3b* and *Dnmt1* in human ESCs results in rapid cell death (Liao *et al.*, 2015).

Active demethylation is orchestrated by members of the Ten-eleven translocation (Tet) family of methylcytosine dioxygenases Tet1, Tet2 and Tet3 (Tahiliani *et al.*, 2009; for a recent review, see Kohli and Zhang, 2013). *Tet1* and *Tet2* are highly expressed in ESCs whereas *Tet3* is induced upon differentiation. Tet1 knockdown in pre-implantation embryos results in a bias to TE formation, highlighting Tet1's role in early embryonic lineage decisions (Ito *et al.*, 2010). Similarly, *Tet1* mutant mouse ESCs display altered differentiation potential and skewing to mesendodermal and TE lineages (Koh *et al.*, 2011). Simultaneous deletion of all Tet genes causes promoter hypermethylation, deregulation of developmental genes and restricts developmental potential of the mouse ESCs (Dawlaty *et al.*, 2014).

Genome-wide mapping of ^{5m}C and ^{5hm}C in ground-state ESCs in 2i ESCs, serum ESCs and EpiSCs revealed characteristic methylation/hydroxylation patterns for each cell state. Such ground-state

ESCs exhibit an altered distribution of ^{5m}C and ^{5hm}C and lower absolute levels relative to serum ESCs. EpiSCs - which are more primed - exhibit increased promoter ^{5m}C coupled with reduced ^{5hm}C, most probably contributing to their reduced developmental potential. Remarkably, a switch to 2i triggers rapid onset of both the ground-state gene expression program and global DNA demethylation. Repression of Dnmt genes by Prdm14 drives DNA demethylation at slow kinetics, whereas Tet1/Tet2-mediated ^{5hm}C conversion enhances both the rate and extent of hypomethylation. These enzymatic systems thus act synergistically during transition to ground-state pluripotency and promote a robust, hypomethylated state (see Fig.4; Hackett *et al.*, 2013)

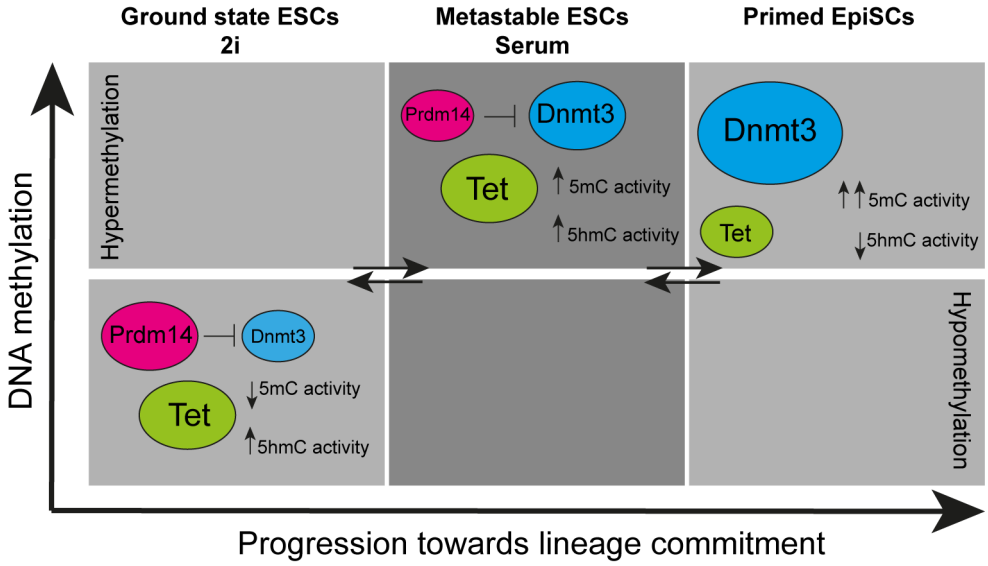

Figure 5. Schematic representation of Dnmt and Tet levels with ^{5m}C and ^{5hm}C state and "serum ESCs" and in EpiSCs.

Ground-state ESCs exhibit high ^{5hm}C levels and Prdm14 activity that represses Dnmt genes. EpiSCs exhibit increased promoter ^{5m}C coupled with reduced ^{5hm}C and high Dnmt expression. ESCs in serum conditions are metastable (Figure modified from: Hackett *et al.*, 2013).

1.2.6.2 Histone signatures

Polycomb group proteins (PcG) modify chromatin structure by (i) depositing methyl marks on H3K27, which leads to repression of target gene expression (Cao *et al.*, 2002) and (ii) mono-ubiquitination of H2A (H2AK119ub1) (reviewed in Campos and Reinberg, 2009). There are two types of Polycomb repressive complex: PRC1 (encompassing canonical and non-canonical PRC) and PRC2.

The **canonical PRC1 complex** is composed of Cbx, Pcgf, Ring and Phc proteins. Cbx proteins recognize and bind H3K27 and facilitate recruitment of PRC2 to the target genes. **The noncanonical PRC1 complex** is composed of Rybp/Yaf2, Ring1a/b, and one of Nspc1/Pcgf1,

Mel18/Pcgf2 (or Mblr/Pcgf6), and it can function more independently - for example - to repress germline-specific genes (Turner and Bracken, 2013).

The **PRC2 complex** - composed of the core components Ezh2, Eed, and Suz12 - mediates trimethylation of H3K27 (H3K27me3). Ezh2 is one of the core components of the PRC2 complex as it catalyzes the H3K27 trimethylation reaction. *Ezh2 null* embryos die before the end of gastrulation (O'Carroll *et al.*, 2001). *Ezh* KO ESCs can be maintained *in vitro* but show defects in differentiation, especially towards mesoderm (Shen *et al.*, 2008). *Eed* KO embryos fail to develop beyond the gastrulation stage (Faust *et al.*, 1995) and *Eed* KO ESCs show even more severe differentiation defects than *Ezh2* KO cells (Shen *et al.*, 2008). Suz12 is essential for PRC2 activity and its gene inactivation results in early lethality (E8.5) of mouse embryos (Pasini *et al.*, 2004). *Suz12* KO ESCs are characterized by global loss of H3K27me3, higher expression levels of differentiation genes and lack of proper downregulation of pluripotency genes (Pasini *et al.*, 2007). The PRC1 complex mediates compaction of the chromatin marked by H3K27Me3. Ring1 is the key mediator of this process (Eskeland *et al.*, 2010). In absence of Ring1b the expression of other genes encoding PRC1 components is reduced, which ultimately leads to derepression of lineage gene expression in ESCs (Leeb and Wutz, 2007). It is also important to note that PRC1 and PRC2 are also crucial for establishing and maintaining long-range interactions during transition from ground to primed state of pluripotency in ESCs (See section 1.2.7 and Joshi et al., 2015; Schoenfelder et al., 2015)

Chromatin where both repressing and activating marks are present (*e.g.* H3K27me3 and H3K4me3, respectively) is referred to as bivalent. In serum ESCs bivalent chromatin marks can be found on key developmental genes where it is thought to keep these genes silenced, but yet poised for activation (Bernstein *et al.*, 2006). Interestingly, prevalence of H3K27me3 and bivalent marks are reduced in 2i ESCs (Marks *et al.*, 2014). H3K4me3 is catalyzed by Trithorax group (TrxG) proteins and is associated with transcriptional activation. TrxG proteins display opposing functions to PcG proteins during development (Mikkelsen *et al.*, 2007; Ringrose and Paro, 2004). Yet, little is known about the role of TrxG protein complexes in ESCs. Wdr5 (one of the key members of the TrxG group) regulates ESC self-renewal through direct protein-protein interaction with Oct4, and Wdr5 knockdown leads to a similar phenotype as found in *Nanog* inactivation or LIF depletion (Ang et al., 2011).

Histone acetyltransferases (HATs) and histone deacetylases (HDACs) control the acetylation status of histones. HATs add acetyl groups to the histone tails, neutralizing them and weakening their nucleosome interactions. HDACs remove acetyl groups from histones and drive chromatin compaction and gene silencing (reviewed in Yang and Seto). ESC chromatin is open (Gaspar-Maia *et al.*, 2011). The histone acetyltransferase Mof plays a key role in the regulation of the core transcriptional network in ESCs. Predominantly it binds together with Nanog to its target genes (79% of Mof-binding sites were

also bound by Nanog). *Mof* deletion leads to loss of self-renewal and defects in embryoid body (EB) formation. Interestingly, overproduction of Nanog in Mof-deficient ESCs partially rescues the self-renewal defects, suggesting that Nanog acts downstream of Mof in ESCs (Li *et al.*, 2012a).

HDACs comprise a large group of enzymes (HDAC1 to HDAC11 and the sirtuin family members SIRT1 to SIRT7). They remove acetyl groups from the histones, thereby increasing their overall positive charge, and allowing closer interaction with DNA. They are very often part of large chromatin remodeling complexes such as NuRD, Sin3 and co-REST (see below and reviewed in Delcuve *et al.*, 2012). *Hdac1* and *Hdac2* KO mouse embryos die perinatally due to multiple defects including in the heart (Montgomery *et al.*, 2007).

Recently, it was shown that Sirt6 controls stem cell commitment via regulation of Tet1-dependent DNA hydroxylation. In absence of *Sirt6*, these KO ESCs subjected to a general differentiation protocol show skewing towards neuroectoderm, do not properly downregulate the pluripotency genes, increase Tet1 expression and - as a consequence - accumulate ^{5hm}C marks. We mention this here as we will report here on similar observations in our Zeb2 research. Sirt6 directly represses *Oct4* and *Sox2*. The authors hypothesize that absence of Sirt6 leads to derepression of these factors that maintain high Tet1 levels, which leads to an increase of ^{5hm}C on neural gene promoters and induction of neural gene expression. Knockdown of Tet1 in the *Sirt6* KO ESC rescues the *Sirt6* KO phenotype. This study elegantly links histone deacetylation with DNA-methylation and cell fate (Etchegaray *et al.*, 2015).

1.2.6.3 ATP-dependent chromatin remodeling complexes

The chromatin state can be altered by chromatin remodeling complexes, which change the histone-DNA contacts, and use the released energy of ATP hydrolysis for this. They can be divided into the SWI/SNF (switch/sucrose non-fermentable), CHD (chromodomain helicase DNA-binding) and ISWI (imitation switch) complexes (reviewed in de la Serna *et al.*, 2006).

BRG- or BRM-associated factor (BAF) and polybromo BAF (PBAF) are the two main complexes in the **SWI/SNF family** (Moshkin *et al.*, 2007). BRG1 (also known as SMARCA4) is the catalytic subunit of the BAF complex in ESCs (EsBAF). It is downregulated upon ESC differentiation. Knockdown of Brg1 results in decreased proliferation rate and reduced differentiation capacity of such ESCs (Ho *et al.*, 2009). Moreover, the BAF complex was shown to occupy the same regions as the key

pluripotency factors suggesting that this complex is involved in the regulation of stemness (Kidder *et al.*, 2009).

The **CHD family** consists of Chd1, 2, 7 and 8 as well as the Nucleosome remodeling and deacetylation complex NuRD. Chd1 is required for the ESCs to maintain their open chromatin state, and knockdown of Chd1 leads to accumulation of heterochromatin (compacted chromatin). Moreover, Chd1 binding strongly correlates with that of H3K4me3 (active mark) whereas bivalent domains are largely devoid of Chd1. Knockdown of Chd1 impairs ESC self-renewal, leads to downregulation of *Oct4* and to an increase in neural gene expression in pluripotency-supporting conditions. Also, when subjected to general EB differentiation, Chd1 knockdown ESCs displayed defects in mesodermal and endodermal differentiation and such cells were more prone to acquire a neural fate (Gaspar-Maia *et al.*, 2009).

The NuRD complex consists of HDAC1 and HDAC2, the histone-binding proteins RbAp46 and RbAp48, the metastasis-associated proteins MTA1 (or MTA2 / MTA3), the methyl-CpG-binding domain protein MBD3 (or MBD2) and the ATP-dependent CHD3 (also named Mi-2alpha) or CHD4 (Mi-2b) (Xue *et al.*, 1998). The repressive activity of NuRD is required for ESC differentiation. In Mbd3-deficient ESCs the NuRD complex does not assemble. *Mbd3* KO ESCs continue to self-renew in absence of LIF and have the capacity to initiate early differentiation, but such cells fail to lineage-commit (Kaji *et al.*, 2006). The NuRD-mediated repressive effect produces transcriptional heterogeneity, also observed in ESCs cultured in serum and LIF conditions; NuRD directly controls this transition out of the self-renewal state by enabling cells to silence expression of a number of pluripotency-associated genes including *Klf4*, *Klf5*, *Tbx3* and *Zfp42* (Reynolds *et al.*, 2012).

The **ISWI family** chromatin remodeling complex consist of three distinct complexes: nucleosome-remodeling factor (NURF), chromatin accessibility complex (CHRAC) and ATP-utilizing chromatin assembly and remodeling factor (ACF), but the role of these complexes in not yet well studied in the context of ESCs. Bptf is the largest subunit of the NURF complex. *Bptf* KO embryos die around E8.0 due to multiple defects and *Bptf* KO ESCs subjected to general differentiation display abnormal meso-, endo- and ecto-dermal development (Landry *et al.*, 2008).

Examples of knockout and/or knockdown phenotypes of selected epigenetic modifiers			
		Mouse KO and/or KD phenotypes	**Mouse KO and/or KD phenotypes**
DNA-methylation	**Dnmt3a**	KO: die at 1 month; (Okano et al., 1999)	Triple KO in mESCs does not affect self-renewal; (Tsumura et al., 2006) Triple KO in hESCs results in rapid cell death; (Liao et al., 2015)
	Dnmt3b	KO: perinatal death; (Li et al., 1992)	
	Dnmt1	KO: perinatal death; (Li et al., 1992)	
DNA-hydroxylation	**Tet1**	KD in preimplantantation embryos results in bias to trophoectoderm formation; (Ito et al., 2010) Depending on the laboratory where the Tet1 KO mice were made, the phenotype can differ: 1) Tet1 (paternal) knockout mice exhibit various phenotypes, including fetal and postnatal growth defects, and neonatal and embryonic lethality; (Kohli and Zhang, 2013) 2) Tet1 mutant mice are viable, fertile and grossly normal though some mutant mice have a slightly smaller body size at birth; (Dawlaty et al., 2011)	KD results in skewing towards mesendoderm and trophoectoderm; (Koh et al., 2011)
	Tet2	KO leads to myeloid malignancies; (Li et al., 2011)	Simultaneous KO of Tet1,Tet2 and Tet3 results in restricted developmental potentialand gene promoter hypermethylation; (Dawlaty et al., 2014)
	Tet3	In Tet3-deficient zygotes, paternal-genome conversion of 5mC into 5hmC fails to occur. Deficiency of Tet3 also delays activation of a paternally derived Oct4 transgene in early embryos Oocytes lacking Tet3 have a reduced ability to reprogram the injected nuclei from somatic cells; (Gu et al., 2011)	
Polycomb Repressor Complex 2	**Ezh2**	KO embryos die before the end of gastrulation; (O'Carroll et al., 2001)	KO show differentiation defects, especially towards mesoderm; (Shen et al., 2008)
	Eed	KO embryos fail to develop beyond gastrulation stage; (Faust et al., 1995)	KO show severe differentiation defects; (Shen et al., 2008)
	Suz12	KO is embryonic lethal (E8.5); (Pasini et al., 2004)	KO display global loss of H3K27me3, higher expression levels of differentiation genes and lack of proper downregulation of pluripotency genes (Pasini et al., 2007).
Histone Acetyltransferases	**Mof**	KO embryos fail to develop beyond the expanded blastocyst stage and die at the implantation stage; (Thomas et al., 2008)	KO leads to loss of self-renewal and defects in embryoid body (EB) formation; (Li et al., 2012a)
Histone Deacetylases	**HDAC1**	die perinatally due to multiple defects including in the heart; (Montgomery et al., 2007)	Double HDAC1/HDAC2 KO results in loss of cell viability, which is associated with chromosome segregation defects; (Jamaladdin et al., 2014)
	HDAC2	die perinatally due to multiple defects including in the heart; (Montgomery et al., 2007)	
	Sirt6	KO leads to genomic instability and aging-like phenotype; (Mostoslavsky et al., 2006)	KO show skewing towards neuroectoderm, do not properly downregulate the pluripotency genes, increase Tet1 expression and - as a consequence - accumulate 5hmC marks; (Etchegaray et al., 2015)

ATP-dependent chromatin remodeling complexes: SWI/SNF	**BRG1**	KO die perimplantation; (Bultman et al., 2000)	KD of Brg1 results in decreasedcell proliferation and reduced differentiation capacity; (Ho et al., 2009)
ATP-dependent chromatin remodeling complexes: CHD	**Chd1**	KO results in arrest of epiblast development at E5.5-6.5, prior to the onset of gastrulation; (Guzman-Ayala et al., 2015)	KD leads to accumulation of heterochromatin, Chd1 impairs ESC self-renewal, KD cells were more prone to acquire a neural fate; (Gaspar-Maia et al., 2009)
ATP-dependent chromatin remodeling complexes: CHD	**Mbd3**	KO die shortly after implantation; (Hendrich et al., 2001)	KO ESCs are able to self-renew in absence of LIF and have the capacity to initiate early differentiation, but such cells fail to lineage-commit; (Kaji et al., 2006)
ATP-dependent chromatin remodeling complexes: ISWI	**Bptf**	Bptf KO embryos die around E8.0 due to multiple defects; (Landry et al., 2008)	KO display abnormal meso-, endo- and ecto-dermal development; (Landry et al., 2008)

Table 1 Examples of knockout and/or knockdown phenotypes of selected epigenetic modifiers

1.2.7 Long-range interactions

Since it was shown that transcriptional regulation of β-globin depends on clustering by proximity of regulatory elements that are key to creating and maintaining active chromatin domains, looping was acknowledged as one of the mechanisms of gene expression regulation (Tolhuis *et al.*, 2002). Dixon *et al.* demonstrated the existence of large "topology associated domains" (TADs) within the genome. These TADs, which are on average 880 kb in size, make-up fundamental structures of mammalian genomes (Dixon *et al.*, 2012).

Kagey *et al.* showed that Mediator and Cohesin physically and functionally connect gene expression and chromatin architecture in ESCs. Mediator complex functions as transcriptional co-activator while cohesin protein complex is a key partner of CTCF and regulator of chromatin looping. shRNA-mediated knockdown of Mediator, Cohesin or Nipbl (Cohesin-loading factor) resulted in phenotypes resembling the knockdown of key pluripotency genes. ChIP-seq analysis of Med1 and Med12 (both subunits of the Mediator complex) showed that at least 60% of actively transcribed genes in ESCs - including enhancer and promoter regions of *Nanog*, *Oct4* and *Sox2* - were bound by Mediator. ChIP-seq for Smc1a and Smc3 (subunits of the Cohesin complex) showed that Cohesin occupies target sites together with or independently of CTCF and Mediator. Interestingly, sites occupied by only Cohesin and Mediator were also associated with RNAPol II (marking active promoters in ESCs). These sites were also bound by Nipbl. Mediator and Cohesin were shown to partner-up to create distinct DNA looping patterns between the gene enhancer(s) and promoter; these loops are cell-type specific and linked to tissue-specific gene expression programs within the cells (Kagey *et al.*, 2010).

The chromatin conformation at the *Nanog* locus in pluripotent cells is dependent on Oct4 presence. Analysis of a 160 kb-long region on chr6 (including *Gdf3*, *Dppa3* and *Nanog*) showed that DNAse-I hypersensitive sites (HSs) and potential regulatory elements are present in this region. Upon Oct4 depletion, the 3D organization of the chromatin changed significantly and the contacts between the HSs within the 160kb-region were greatly diminished (Levasseur *et al.*, 2008). Recently, it was shown that two transcriptional co-activators, p300 and CPB, also play an important role in ESC chromatin looping. P300 and CBP are recruited to the *Nanog* locus to mediate the formation of long-range chromatin looping structures with regions showing enhancer activities (Fang *et al.*, 2014).

Recently, Joshi and colleagues used high-resolution chromosome conformation capture technique (Hi-C) to study spatiotemporal changes in long range interactions during transition from ground to primed state in mESCs and *vice versa* (Joshi et al., 2015). They showed that "extremely long-range promoter-promoter interactions" (ELRIs, defined as long-range interactions - intra and interchromosomal – but excluding intra-TAD interactions) are present exclusively in serum ESCs and that their formation depends on the action of the Polycomb repressive complex PRC2 and is associated with PRC1 activity. Deposition of H3K27me3, mediated by the PRC2 complex, correlated with ELRIs and overlapped at the promoters of genes involved in cell fate determination like the Hox gene cluster indicating that lineage priming in serum conditions involves formation of ELRIs. Notably, ELRIs gradually disappeared upon changing ESC culture medium to 2i. To validate the role of PRC2 complex in the formation of ELRIs, the authors used *Eed* KO ESCs. In these cells the PCR2 complex does not form what leads to absence of H3K27me3 marks. Hi-C evaluation of *Eed* KO ESCs in serum revealed that ELRIs did not form properly and chromosome conformation resembled that of 2i. Moreover, the authors, based on the dynamics of the PRC1 complex binding to H3K27, speculated that it might be involved in fine-tuning and mediation of ELRIs formation.

Another study by Schoenfelder and co-workers, published around the same time as the study by Joshi et al. provided supportive evidence for the role of Polycomb complexes as important regulators of chromosome conformation (Schoenfelder et al., 2015). Using Ring1A-Ring1B-double (d)KO mESCs, the authors showed that ELRIs cannot be established because these dKO mESCs cannot form functional PRC1 complex. Together, these two studies show that PCR1 and PRC2 complexes are master regulators of three-dimensional interaction networks in the genome of mESCs.

In summary, the 3D architecture of the genome and its rearrangements influence transcriptional regulation at various cellular states. Chromatin structures include highly conserved, stable TADs as well as cell-type specific, but dynamic loops. Dynamics of chromatin looping is not only dictated by the actions of key proteins (such as the aforementioned Cohesin, CTCF, Mediator) but can be influenced by various TFs (including pioneer TFs) and chromatin-modifying complexes including PCR1 and PRC2.

1.3 Neurogenesis

1.3.1 In vivo

1.3.1.1 Early neural development (E7.5-E10.0)

The epiblast of the embryo acquires neural fate as default in the controlled inactivity of Wnt, Nodal and Bmp signaling, which otherwise act as mesoderm-inducing and endoderm-permissive signals (Stern, 2005). The activity controlling Bmp antagonists Chordin, Noggin and Follistatin, and the Wnt inhibitor Dkk1 (of the Dickkopf family), and the Nodal inhibitors Lefty and Cerberus (the latter also binding Bmp and Wnt) are key to epiblast cells acquiring neural fate. Neural induction in the mouse embryo starts around E7.5 and rapidly produces a single-layered sheet of cells (the neural plate) that acquires progressively an anterior-posterior identity. The neural plate undergoes reshaping (neurulation) and ultimate closes to make the neural tube. Following the closure, neural progenitors that start to proliferate at the rostral side of the embryo causes the future telencephalon to "balloon out" and generate two brain hemispheres. The remaining neural tube in the cranial region undergoes regional specification and also gives rise to the midbrain and hindbrain (Smith, 1997).

Two signaling centers pattern the early central nervous system (CNS): the anterior neural ridge (ANR), which is responsible for forebrain patterning, and the midbrain-hindbrain border (MHB, also named the isthmus or the isthmic organizer), which regulates early midbrain and hindbrain specification (reviewed in Wurst and Bally-Cuif, 2001). Fgf8, secreted by both ANR and MHB is an important regulator of early CNS development (Martinez *et al.*, 1999). Retinoic acid (RA) is important for anterior-posterior CNS patterning; it is well known for its co-regulation of Hox gene expression and is required for hindbrain and spinal cord development (reviewed in Maden, 2002). Further development of the telencephalon and dorso-ventral patterning is orchestrated by opposing actions of Fgf8 and Bmp (secreted by the roof plate) and Shh (secreted by the ventrally located prechordal plate and in the neural tube by the floor plate cells, as well as by the underlying notochord) (Hoch *et al.*, 2009).

1.3.1.2 Corticogenesis

The early embryonic brain can be subdivided into three parts: forebrain (prosencephalon), midbrain (mesencephalon) and hindbrain (rhombencephalon) (see Fig.6A). The forebrain later consists of the more anteriorly located telencephalon and the posteriorly located diencephalon. The ventral part of the

telencephalon (subpallium) contains ganglionic eminences that give rise to basal ganglia, whereas the dorsal part of the telencephalon (pallium) consists of the (neo)cortex ((neo)pallium), hippocampus (archipallium) and the olfactory lobe (paleopallium) (see Fig. 6A, right panel).

The (neo)cortex is considered to be the "latest evolutionary addition" to the brain structure of mammals in comparison with other vertebrates like *Xenopus*. It first appeared in the small mammalian species during the transition from Triassic to Jurassic period (±200 million years ago). Since, it has undergone significant enlargement during evolution (Finlay and Darlington, 1995) and is considered to be responsible for the mental prowess and extraordinary cognitive abilities of the human species (Rakic, 2009). The mammalian cortex is a six-layered structure composed of possibly hundreds of different types of neuron and glial cell (reviewed in Franco and Müller, 2013). Neurons of the cortex in the forebrain can be subdivided into two main classes: projection (pyramidal, glutamatergic, excitatory cells) that extend their axons not only within the cortex but also to other distant regions within the brain and spinal cord and into local circuit neurons (cortical interneurons, GABAergic, inhibitory cells), which make local synapses. Proper brain function requires balance between excitation and inhibition that is achieved by excitatory and inhibitory neurotransmitters, respectively. Glutamate (produced by the projection neurons) is the main excitatory and GABA (produced by the interneurons) is indeed the main inhibitory neurotransmitter in the mammalian cortex. Disproportions in numbers of excitatory versus inhibitory neurons nearly always lead to seizures and epilepsy (Powell *et al.*, 2003).

Corticogenesis in the mouse starts with the neural tube closure around E10 when a single layer of fast-dividing neuroepithelial progenitors, the ventricular zone (VZ) is formed within the thin cortex. Neuroepithelial progenitors can undergo symmetric cell division generating two identical daughter cells (i.e. two neuroepithelial progenitor cells) or divide asymmetrically to generate one neuroepithelial progenitor cell and one postmitotic neuron, which around E11 migrates out from the VZ to the pial surface to form the preplate (PP, see grey cells in Fig. 6). The second wave of neurogenesis takes place around E13 when the newly formed projection neurons split the PP into a superficial marginal zone (MZ, which contains the first-born Cajal-Retzius neuronal cells, also called layer I of the cortex) and the subplate (SP) and thereby create the cortical plate in-between (CP, see blue cells in Fig. 6; for a recent review on SP neurons, see: Hoerder-Suabedissen and Molnar, 2015). At this time point (E14) the CP contains only one layer of neurons (layer 6), but at the end of neural development it will consist of five layers of distinct subtypes of neurons (numbered 2-6 with layer 6 being the deepest and first formed neuronal layer of the cortex) that originate from a common precursor cell called radial glia (characterized by the presence of GFAP, GLAST and BLBP). Between E14 and E18 layers 2-5 are subsequently formed as the new waves of post-mitotic neurons leave the VZ, migrate radially out and position themselves on top of previously formed neurons (see Fig. 6B). This process is also referred to as the "inside-out" model of neocorticogenesis, is accompanied by visible reduction of the VZ thickness

(reviewed in Gupta *et al.*, 2002) and involves lineage determination of neuroepithelial cells, radial glial cells and neurons. Importantly, the temporal waves of neurogenesis can be recapitulated and are respected *in vitro*, which has prompted the field to mainly think of intrinsic, transcription factor combination based control of neurogenesis.

Our work *a.o.* in this PhD thesis has however identified, via work on Zeb2 *in vivo* in corticogenesis, the existence of extrinsic control achieved via non-autonomous actions of Zeb2, thereby providing a feedback mechanism from upper layer neurons to the VZ (see Chapter 3 of this thesis). More recently, basal progenitor cells able to divide away from the VZ have been identified as cellular intermediates in the generation of neurons via symmetrical neurogenic division (Molyneaux *et al.*, 2007) In addition to the principles summarized above also polarized organization and interkinetic nuclear migration of neuroepithelial cells, radial glial cells and basal progenitors in the cortex show interesting differences (see also: Fietz and Huttner, 2011; Molyneaux *et al.*, 2007).

Each of the six cortical layers contains neurons with distinct molecular signatures and connectivity. Different neuronal subtypes can be characterized by sets of markers that are specific for one or more layers (one or more). For example, layer 1 contain Reelin and Lhx5, layer 2/3 and 5 are marked by Brn2, and Ctip2 is an often used characteristic marker of the deep layers 5 and 6 (for more markers, see Molyneaux *et al.*, 2007). Substantial efforts are directed towards obtaining a full set of markers that can define all subtypes of neuron, or for some their partial overlap in gene expression profile, in the brain based on their molecular (mainly mRNA expression) signatures (Cahoy *et al.*, 2008; Sugino *et al.*, 2006; Wichterle *et al.*, 2013).

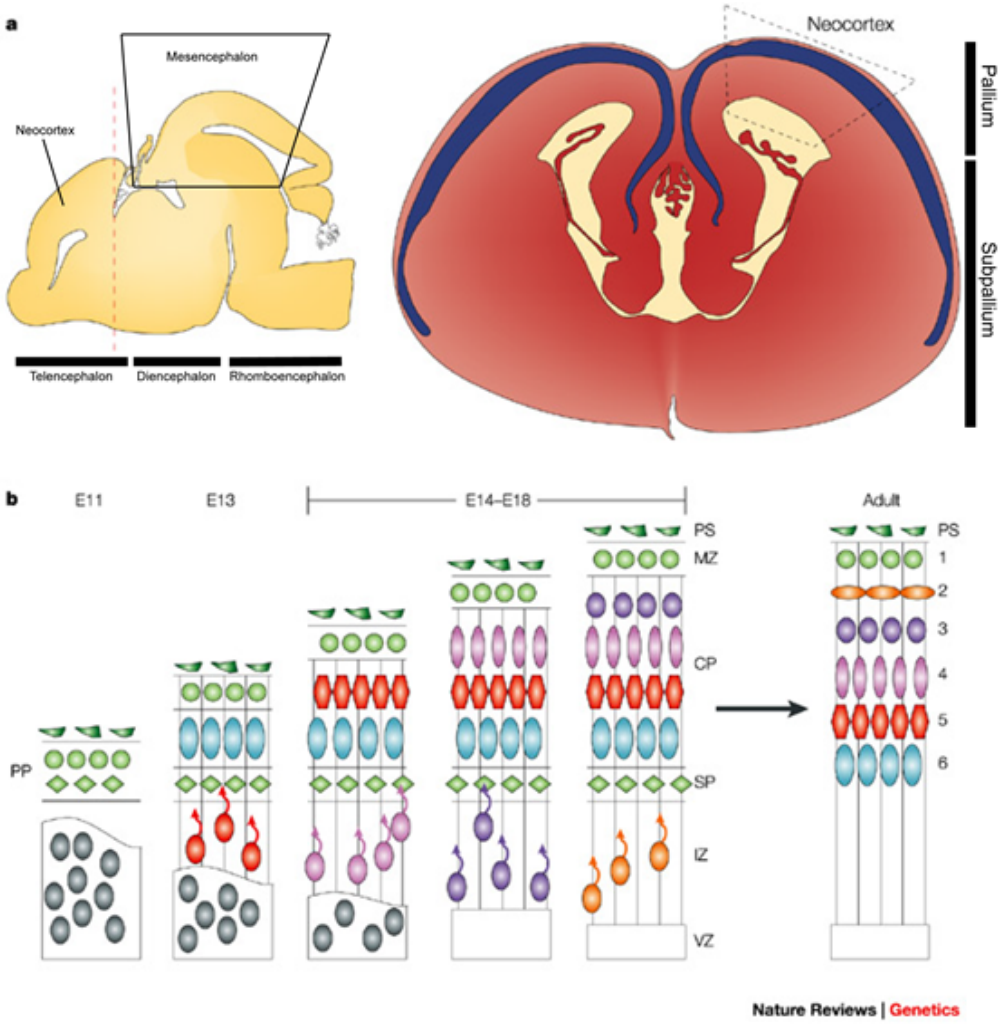

Figure 6. The location of cortex in the central nervous system.

A lateral view of an E16 mouse brain (left). The red dashed line indicates the plane of section from which a coronal section (right) has been taken, which shows the location of the cortex. b. The organization of the adult cortex into distinct neuronal layers. Key developmental stages (E: embryonic day) of the radial component of neocortical-layer formation are shown.
Abbreviations: VZ: ventricular zone, IZ: intermediate zone, SP: subplate, CP: cortical plate, MZ: marginal zone, PS: pial surface (Gupta *et al.*, 2002)

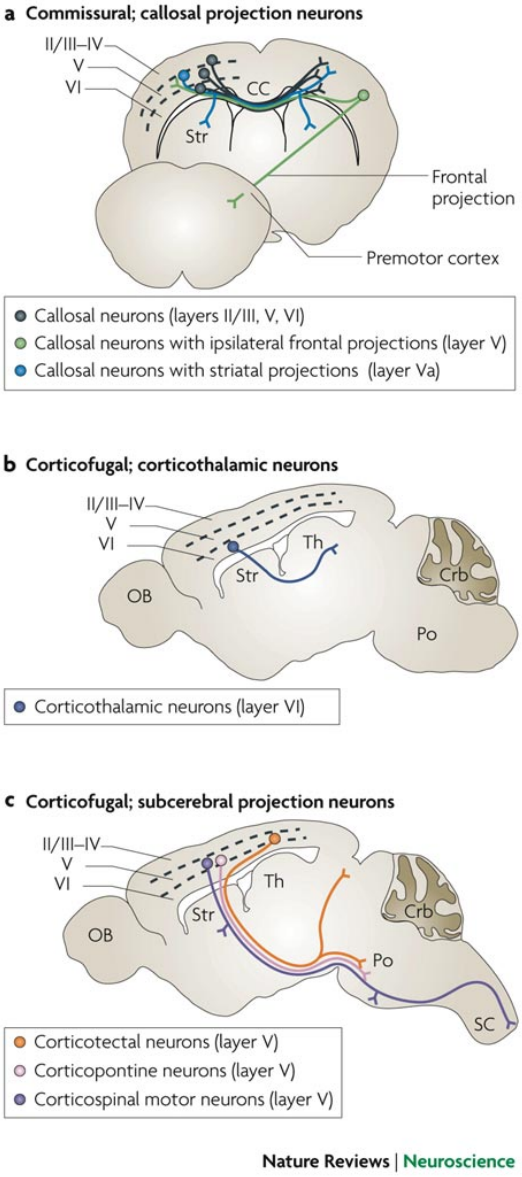

Figure 7. Three basic classes of cortical projection neurons: associative, commissural and corticofugal.

a. Commissural; Callosal projection neurons. Projection neurons of small to medium pyramidal size that are primarily located in layers II/III, V and VI, and extend an axon across the corpus callosum (CC) (panel a). At least three major types of callosal neuron can be classified. These maintain: single projections to the contralateral cortex (black); dual projections to the contralateral cortex and ipsilateral or contralateral striatum (blue); and dual projections to the contralateral cortex and ipsilateral frontal cortex (green). These never project axons to targets outside the telencephalon. Str, striatum. b. Corticofugal (subcortical); Corticothalamic neurons. Projection neurons primarily located in cortical layer VI, with a smaller population in layer V, that project subcortically to different nuclei of the thalamus (Th) (panel b). c. Subcerebral projection neurons; Also referred to as type I layer V projection neurons (panel c). These include pyramidal neurons of the largest size, which are located in deep-layer V and extend projections to the brainstem and spinal cord. They can be even further subdivided into several distinct projection neuron subtypes. Amongst them: Corticotectal neurons (orange), Corticopontine neurons (pink), Corticospinal motor neurons (purple).
Other abbreviations: Crb, cerebellum; OB, olfactory bulb; SC, spinal cord. (Figure and legend taken from: Molyneaux *et al.*, 2007)

Based on the combination of molecular signature with connectivity, there are grossly three subclasses of projection neuron. *i.e.* associative projection neurons (that extend axonal projections within one hemisphere of the brain), commissural projection neurons (that project to the opposite hemisphere via the corpus callosum, *i.e.* the callosal neurons or the anterior commissure) and corticofugal projection neurons (that project outside of the cortex, referred to as subcerebral and corticothalamic projections). Commissural neurons primarily occupy layers 2/3 (for about 80% of these cells), 5 and 6 (20%). Corticofugal neurons are located in layer 6 and 5, while associative neurons can be found in all layers of the cortex (Fig.7) (reviewed in Custo Greig *et al.*, 2013).

There are two major types of glial cell in the brain, *i.e.* macroglia and microglia. Astrocytes, oligodendrocytes and ependymal cells represent the three differentiated macroglial cell types. They are derived from precursor cells located in the VZ of the developing cortex. Astrocytes are star-shaped cells that provide support to other brain cells, regulate calcium influx and synaptic transmission, and maintain the brain-blood barrier (reviewed in Molofsky *et al.*, 2012). Oligodendrocytes provide insulation for the axons, which is critical for rapid nerve impulse conduction (Emery, 2010). Ependymal cells are derived from radial glial precursors (the 4th glial cell type present in the developing brain, giving rise to neurons and glia in the brain, see above) and they play major role in the transport of cerebrospinal fluid and brain homeostasis (Spassky *et al.*, 2005). Microglia originate from hematopoietic stem cells and function as macrophage-like cells to immune-protect the CNS (Kettenmann *et al.*, 2011); more recently additional roles of distinct microglia cells have been identified to support adult neurogenesis (Ribeiro Xavier *et al.*, 2015) and neural precursor cells in general (Su *et al.*, 2014).

Forebrain interneurons are born in a progenitor cell zone in the ventral telencephalon. The ventral telencephalon can be divided into four major parts - medial ganglionic eminence (MGE), lateral ganglionic eminence (LGE), caudal ganglionic eminence (CGE) and preoptic area (POA) - each of which produces specific (identifiable) subtypes of interneuron that populate the cortex during development (see Fig. 8).

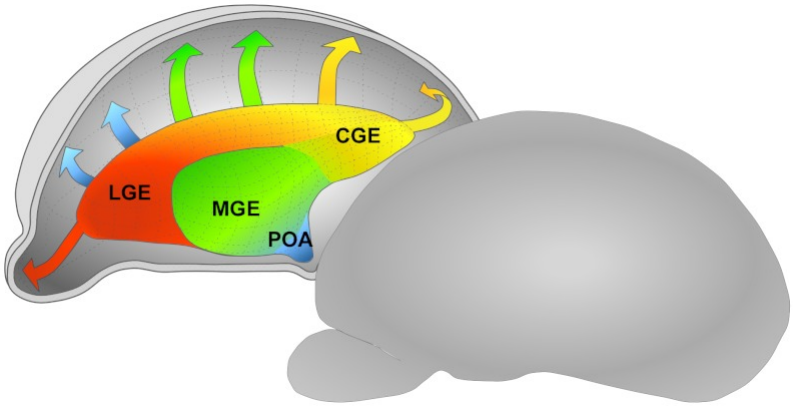

Figure 8. Cortical interneurons are born in the subpallium and migrate tangentially to the cortex.

The scheme represents an E13.5 embryo brain hemi-section. The arrows show representative migratory routes. POA-derived interneurons (blue) have a bias to invade the cortex through its rostral region, while CGE-derived interneurons (yellow) primarily reach the cortex by its caudal pole. The LGE (red) is a major source for many olfactory bulb interneurons, although some of the latter are also produced in other regions. For easier representation, the septum and the thalamus are not depicted in the schema.

Cortical interneuron diversity emerges from distinct progenitors pools that are spatially segregated within the VT (via the respective MGE, LGE, CGE and POA pools). Approximately 50-60% of all cortical interneurons originate from the MGE, 30-40% originate from the CGE, a small fraction of cortical interneurons emerges from POA, whereas LGE is the primary source of the olfactory bulb interneurons see Fig. 9) and an important source for stem/progenitor cells of the subependymal zone (SEZ, now often designated as SVZ) of the lateral ventricles in the adult forebrain, an acknowledged stem cell compartment (Furutachi *et al.*, 2015; Lin and Iacovitti, 2015; Urbán and Guillemot, 2014), where Zeb2 also plays a role (Stappers *et al.*, *unpublished results*).

All interneurons follow distinct routes of tangential migration to reach their final destination within the brain (for a simplified schematic representation of the interneuron routes, see Fig. 9) in a process that starts around E12.5 and lasts till birth. Because of the enormous diversity of the interneuron types (considering their birth place and the precursor cell type, final destination in the brain, morphology, connectivity, molecular signature and electrophysiological properties) there is no comprehensive list/scheme yet describing all subtypes of interneurons present in the brain. The most commonly used and somewhat old system to distinguish between the different interneuron subtypes is (still) based on the presence of calcium binding proteins such as parvalbumin (PV), calbindin (CB) or calretinin (CR) and neuropeptides such as somatostatin (SST), vasoactive intestinal peptide (VIP), neuropeptide Y (NPY), or cholecystokinin (CCK) (Flames and Marín, 2005), in addition to their origin (Xu *et al.*, 2004) and their mode of firing (Gelman, Marín, 2012). Obviously, any dysfunction in

generation, specification, migration, maturation and/or function of cortical pyramidal neurons and/or interneurons and/or glia is likely to contribute to severe neurodevelopmental disorders.

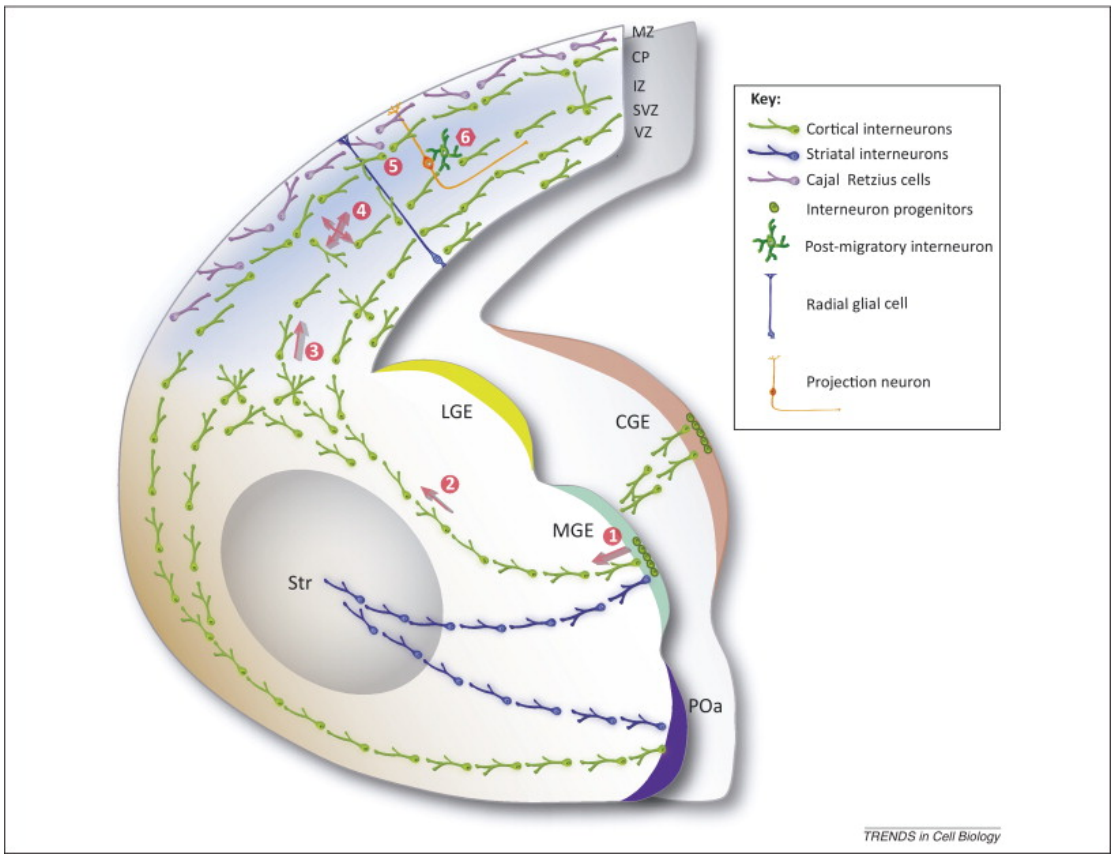

Figure 9. Patterns of interneuron migration in the developing telencephalon.

This scheme shows a rostral and caudal hemi-section through the mouse telencephalon at the mid-embryonic (E15) stage. The major decision-making steps (1–6) involved in the migration of cortical interneurons derived from the subpallium are illustrated. Interneurons derived from the medial ganglionic eminence (MGE) (green), the preoptic area (POa) (purple), or the caudal ganglionic eminence (CGE) (orange) exit the proliferative zones and initiate their migration toward the developing cortex and striatum. Arrows indicate net directionality of movement. *Abbreviations: LGE, lateral ganglionic eminence; Str, striatum; VZ, ventricular zone* (Guo and Anton, 2014).

1.3.2 In vitro

Since the establishment of the first ESC lines, a number of protocols for their neural differentiation have been published, however, it is still a challenge to produce efficiently sufficient pure subtypes of neuron and glia in a dish (reviewed in Cai and Grabel, 2007; Germain *et al.*, 2010). The simplest protocol for neural differentiation (ND) of ESCs was inspired by the *in vivo* data showing that in absence of anti-neurogenic signals cells of the epiblast acquire neural fate (see Chapter 1.3.1.1). This *monolayer*

differentiation protocol, which varied depending on the ESC line used, produces relatively pure neural populations, however without additional signaling activation/inhibition it is not possible to direct the ESCs to differentiate to a desired neuronal subtype (Ying *et al.*, 2003).

Recently, it was shown that by applying small modifications to the Ying protocol (Ying *et al.*, 2003) it is possible to mimic cortical development in a dish. For example, Gaspard *et al.* (2009) showed that ESCs cultured in chemically-defined medium in the absence of morphogens and in the presence of the Shh inhibitor cyclopamine (again, inspired by the *in vivo* data, see Wilson and Rubenstein, 2000) could recapitulate most important steps of neuron development and ultimately produce pyramidal neurons displaying layer-specific identity (Gaspard *et al.*, 2009). Aggregation of ESCs into EBs exposed to neural-inducing activity of RA (based on older experiment that used EC cells) was shown to result in almost pure neural populations even when serum was present in the medium (Bain *et al.*, 1996). In fact, this protocol represents a good alternative for the ESC lines that do not perform well in the monolayer differentiation protocol of Smith and co-workers. One should note however that RA acts in concentration-dependent manner and high RA induces more caudal fate (Okada *et al.*, 2004).

Several protocols where EB formation was combined with signal inhibition were also established. EB formation in medium containing knockout serum replacement (called "serum-free suspension culture", SFEB) and Wnt and Nodal inhibitors (Dkk1 and Lefty1, respectively) induced very efficient ND and production of Bf1-positive telencephalic neurons, which could be further specified by introducing Wnt3a (for pallial) or Shh (for basal) telencephalic specification (Eiraku *et al.*, 2008; Watanabe *et al.*, 2005).

A promising study done by Lancaster *et al.* described derivation of an organoid-type culture system, named "cerebral organoids", obtained from human ESCs, in which development of various discrete brain regions occurred in the course of 30-75 days. This protocol involved EB formation (d0-d4, in human ESC medium) that was followed by transfer to neural induction medium (d4-d11), embedding in matrigel, stationery differentiation (d11-d15) and, finally, transfer to a spinning bioreactor for the rest of the duration of the protocol. Interestingly, the neural induction media did not contain inhibitors of known neural-inhibitory pathways. The cerebral organoids differentiation showed remarkable resemblance to cerebral cortex development. The "mini-brains" were composed of progenitor zone cell populations with abundant radial glia layer and forebrain-specific neurons with highly organized layered structures (Lancaster *et al.*, 2013).

1.4 Zeb2 in development

1.4.1 Discovery of Sip1/Zeb2

Sip1/Zeb2 was identified in a yeast 2-hybrid screening partners for receptor-regulated Smads by using the transcriptionally active Smad1-MH2 domain (Meersseman et al., 1997) as a bait and polypeptides expressed from an in-house made (L. Nelles, unpublished results) E12.5 mouse embryo cDNA library as a prey.

The screening in yeast resulted in ±80 candidate Smad-interacting proteins (SIPs), about half of which have meanwhile in the field been identified by others as well and confirmed as SIPs (Zwijsen et al., 2003; unpublished results). In our laboratory the decision was made to first continue working on the protein encoded by the picked th12 partial cDNA sequence (later renamed Sip1 as full-length coding cDNA sequence, see below), for its coding sequence showed high homology with the previously characterized DNA-binding zinc finger, (non-DNA-binding) homeodomain-like domain containing TF δEF1 (Funahashi et al., 1991; Sekido et al., 1996), later also identified by many other teams and given multiple alternative names, including *a.o.* Zfhx1a and eventually Zeb1). This high similarity between th12 and δEF1 was confirmed for the full-length cDNA coding sequence of Sip1 (Verschueren et al., 1999).

Sip1's interaction with full-length Smads in mammalian cells was shown to depend on receptor-mediated Smad activation, and Sip1 was found to bind to all such Smads tested, i.e. Smad2/3 for TGFβ/activin signalling and Smad1/5 for BMP signalling (later followed by Smad8; Yoshimoto et al., 2005) Meanwhile in our lab, using th12 as a bait, also Smad2 was picked-up in a yeast 2-hybrid screening, together with many other candidate th12-interacting polypeptides (van Grunsven, unpublished results).

Like δEF1, and considering the very high identity between the DNA-binding zinc fingers in the two zinc finger clusters in the proteins, SIP1 was found to bind to 5'-CACCT sequences in different promoters, including the *Xenopus B*ra promoter. Overexpression from *in vitro* made, injected sense RNA of either full-length Sip1 or a N-terminally truncated polypeptide encompassing the C-terminal zinc finger cluster, and which both bind to the Xbra2 promoter *in vitro* in a short region conferring activin responsiveness, prevented expression of the endogenous *Xbra* gene in early *Xenopus* embryos. Therefore, Sip1, like δEF1, was proposed to be a transcriptional repressor, which may be involved in the regulation of at least one immediate response gene in *Xenopus* for Nodal/Activin-dependent signal transduction pathways (Verschueren et al., 1999).

Subsequent work in our lab mapped the DNA-binding mode of full-length members of this small family of vertebrate TFs. From a number of *in vitro* DNA-binding experiments it was proposed that each zinc finger cluster binds independently to a separated repeat of the 5'-CACCT sequence in order to bind the promoter regions of candidate target genes like *Xenopus Xbra2*, human *alpha4-integrin*, *E-cadherin* (later identified as an important Sip1 target in EMT, including in cancer; Comijn et al., 2001) and *follistatin*, hence forming high-affinity bipartite elements composed of one CACCT and (often) one CACCT(G) sequence, the orientation and spacing of which can indeed vary (Remacle et al., 1999). Using transgenic *Xenopus* embryos (collaboration with J. Smith, London), our laboratory then demonstrated that the integrity of these two sequences is necessary for correct spatial-temporal expression of a *Xbra2* promoter-driven reporter gene (Lerchner et al., 2000). We also found that both zinc finger clusters must be intact for the high-affinity binding of Sip1 to DNA and for its optimal repressor activity. Our results showed further that Sip1 likely binds as monomer and contacts one target sequence with the first zinc finger cluster, and the other with the second cluster. Hence, our work redefined the optimal binding site and, consequently, candidate target genes for vertebrate members of the (later renamed as) Zeb family.

1.4.2 Sip1/Zeb2 protein partners

Besides Sip1-Smad interaction (which depends on 4aa in the Sip1 SBD; Conidi et al., 2013), also Sip1 interaction with the co-repressors CtBP1/2 (van Grunsven et al., 2003) and the NuRD complex (Verstappen et al., 2008) respectively, were identified and further documented in our laboratory.

Recently, R. Poot and co-workers (Erasmus MC, Rotterdam) identified Zeb2 as a partner of Sox2 in NSCs (Engelen et al., 2011), later confirmed by A. Conidi in mESCs (unpublished). The overlap between Mowat-Wilson syndrome (See Chapter 1.5.1) and SOX2 anophtalmia syndrome (SAS, caused by SOX2 mutations) and accompanying intellectual disability (ID) may point to one or more pathways wherein Sox2 and Zeb2 co-operate. This is further being investigated and combined with the identification of the Zeb2 (in various cell systems, including NSCs) and Sox2 interactome and of the bHLH factor Tcf4 (E2-2) in NSCs. TCF4 mutations cause Pitt-Hopkins syndrome (PHS), a severe ID syndrome that shares many features with non-mild MOWS, including the complete lack of speech in patients. Intriguingly, among the identified Tcf4 partners were both Sox2 and Zeb2. This suggests a complex between these three factors that each cause ID, albeit with different penetrance and severity.

Currently, in collaboration with the laboratory of R.Poot, our lab is exploring functional relationships between Zeb2, Tcf4 and Sox2 and their significance during cell differentiation and specification.

In addition, using a flag-V5-tagged Zeb2 overexpression construct, our laboratory performed new proteomic screen in neural stem cells (unpublished data). We confirmed interaction of Zeb2 with Tcf4, Tcf12, NuRD (Chd, Mta, Smarc and others), i.e. ATP-dept chromatin remodelers, CtBPs-Wiz-histone methyltransferases, co-repressors Ctbp-1 and -2, SWI-SNF subunit-related proteins and we identified novel interacting partners for Zeb2 that include: Sox2/5/6, many Zfps , Sall proteins, Rcor-REST, Mediator complex proteins, HDACs, Ehmt (H3K9me), Lys-specific demethylases, Protein phosphatase regulatory and catalytic subunits, SUMO binders and activators, Ub binders, Trim proteins, some of which are a E3 ubiquitin-protein ligase, importin-7, a number of RNA-binding proteins, Nono , caseine kinase subunits, Man1-Lemd3, YY1, Menin-1, Tet1 and OGT (See also General Discussion in this PhD thesis).

1.4.3 Neural phenotype of the *Zeb2* conventional knockout mouse

During embryogenesis, the first Zeb2+ cells can be detected from neural plate-late bud (NPLB) stage in the entire mesoderm (of the embryo and in extraembryonic mesodermal tissues) (Fig. 10a). Zeb2+ were first detected in the anterior neural plate starting from E7.5 (Fig. 10b") (with the exception of the most anterior part of the neural plate, see Fig. 10b') and later spread to the posterior part of the neural plate (Fig. 10 c"). Also, Zeb2 was detected in the migrating neural crest cells (Fig. 10 c') and in the somites (Fig. 10 c", PhD thesis Debruyn, 2010).

To understand the role of Zeb2 in early development, a conventional *Zeb2* knockout mouse model was generated using a ubiquitous Cre-*deleter* strain removing the critical exon 7 from a floxed mouse Zeb2 allele (Higashi *et al.*, 2002). Conventional Zeb2-deficient (Higashi *et al.*, 2002) *i.e.* homozygous mutant embryos develop normally until the late headfold stage (E7.5-7.75). The first defects become visible from the onset of somitogenesis (around E8.0-8.5), but also include lack of neural tube closure with reduced expression of *Sox2* and overexpression of E-cadherin mRNA and protein (Van de Putte *et al.*, 2003), defects in the delamination/migration of cranial neural crest cells (which are still formed there) and lack of vagal neural crest cell production (Van de Putte *et al.*, 2003), disappearance of Krox20+ cells in the rhombomeres 3 and 5 of the hindbrain due to excessive apoptosis (Van de Putte 2003, 2007 and *unpublished results*), abnormal development of allantois (Debruyn, 2010) and aberrant morphology of the somites (Maruhashi *et al.*, 2005). Conventional *Zeb2* KO embryos die around E9.5 and show severe growth retardation and never undergo embryonic turning (Higashi et al., 2002; Maruhashi et al., 2005; Van de Putte et al., 2003).

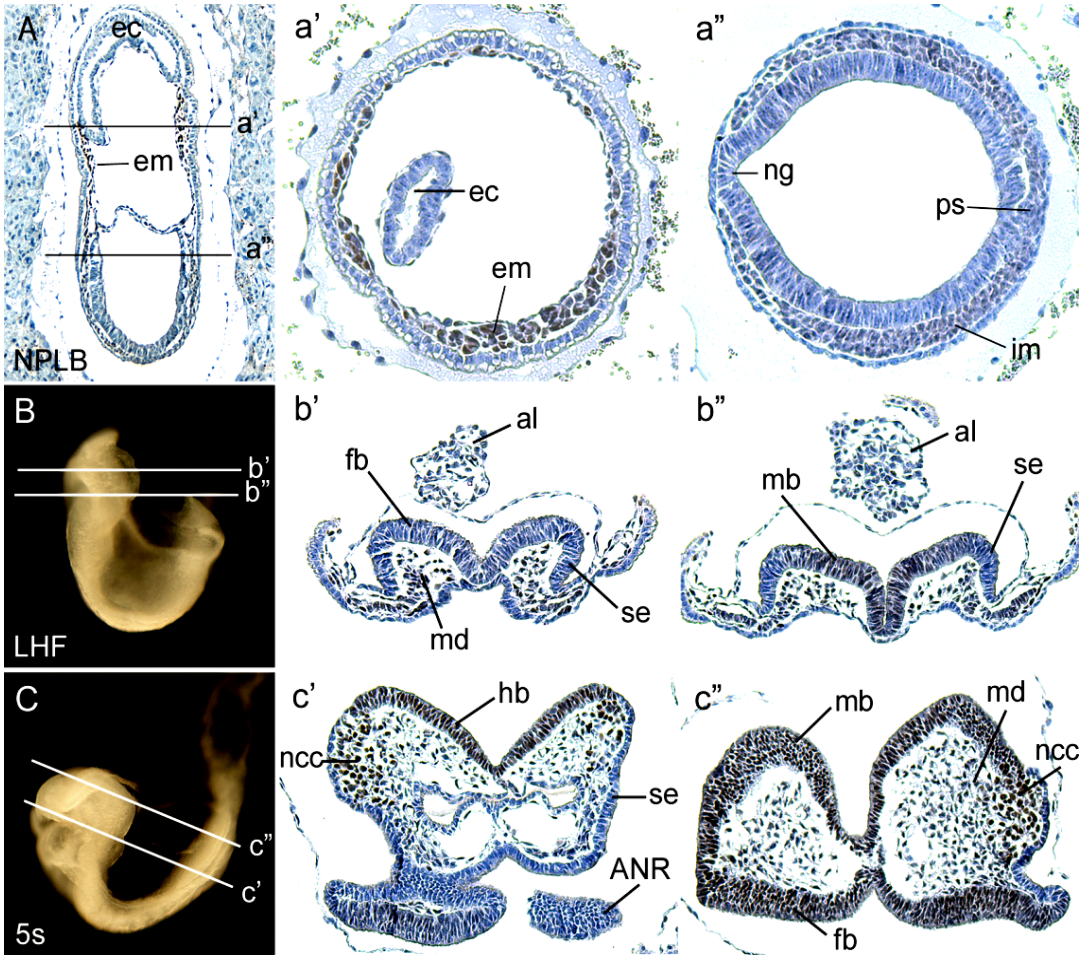

Figure 10. Zeb2/Sip1 is present in the neural plate from the headfold stages onwards.

(A) Sip1 is detected both in the extra-embryonic (a') and the embryonic mesoderm (a") at the NPLB stage. (B) Sip1 is first detected in the neural plate at the headfold stages; no Sip1 protein can be detected in the most anterior forebrain (b') while Sip1 is already clearly present in the rest of the future brain region (b"). (C) At the 5s stage, Sip1 is present in the entire neural plate, but remains absent in the ANR (c'). Sip1 can also be detected in the migratory neural crest cells (c'). Sip1 is never expressed the non-neural surface ectoderm (b", c',c"). (A) Saggital section; (B and C) are lateral views of whole embryos, anterior is towards the left in (A-C). All other sections are transversal, anterior is towards the left in (a' and a"), dorsal is towards the top in (b'-c"). All results are obtained by immunohistochemistry.

Abbreviations: al, allantois; ANR, anterior neural ridge; ec, ectoplacental cavity; em,extra - embryonic mesoderm; fb, forebrain; hb, hindbrain; im, intraembryonic mesoderm; mb, midbrain; md, mesoderm; ncc, neural crest cells; ng, neural groove; ps, primitive streak; se, surface ectoderm. (PhD thesis: Debruyn, 2010)

Later in development and in the adult wild-type mouse *Zeb2* expression becomes restricted to specific cell types especially within the embryonic CNS cortex (Seuntjens *et al.*, 2009), but it can also be detected and plays a role in other parts of the CNS (van den Berghe *et al.*, 2013; McKinsey *et al.*, 2013; Miquelajauregui *et al.*, 2007; Roy *et al.*, 2012), the peripheral nervous system (PNS) (Jeub *et al.*, 2011; Weng *et al.*, 2012) and the enteric nervous system (ENS) (Stanchina *et al.*, 2010), the hematopoietic and innate immune system, and cardiac repair, where our lab – in collaboration with experts in the respective systems – studied also its functional role *in vivo* in the meantime (Goossens *et al.*, 2011) as

well as in certain cancers (Denecker *et al.*, 2014; Goossens *et al.*, 2015), (see section 1.4.2 for a selection).

1.4.4 *Zeb2* conditional, cell-type specific knockouts

Zeb2 was shown to play multiple roles during brain development. Here, we discuss published data addressing the role of Zeb2 in hippocampus development as well as recent studies on the role of Zeb2 in developing interneuron migration and specification. In Chapter 3 we specifically focus on the role of Zeb2 in the development of the projection neurons in the mouse brain cortex and we show that Zeb2 controls both the numbers and timing of the projection neurons and glia generated (see Chapter 3 and Seuntjens *et al.*, 2009).

1.4.4.1 Brain development

Hippocampus is the part of the brain responsible for memory consolidation as well as spatial memory and navigation (Burgess *et al.*, 2002). Our lab – in collaboration with V. Tarabykin (Göttingen, now in Berlin) – showed that Zeb2 is essential for correct hippocampal formation in the developing brain (Miquelajauregui *et al.*, 2007). *Zeb2* is strongly expressed in the hippocampus during development (see also Chapter 3, Fig. 12) and genetic inactivation of *Zeb2* from the cortical progenitors (using the Emx1-Cre deleter mouse) affects hippocampal development from E15.5 onwards. The defects include overall smaller hippocampus and almost absent dentate gyrus and are caused by decreased cell proliferation and increased apoptosis. Since these defects strongly resemble phenotypes resulting from deletion of some of the key components of the Wnt signaling pathway (including knockouts of β-catenin, Frizzeld-9 and Emx2), the expression of the known Wnt pathway inhibitors Sfrp1 and Sfrp2 was examined. The expression of Sfrp2 mRNA was not altered in the *Zeb2* mutants, whereas Sfrp1 mRNA expression was dramatically increased in the post-mitotic cells of the hippocampus of the *Zeb2* knockout brains. Next, direct binding of Zeb2 to the *Sfrp1* promoter was shown. Zeb2 was also shown to specifically regulate the non-canonical (JNK-dependent and β-catenin-independent) signaling in the developing hippocampus. Taken together, this data shows that in the hippocampus Zeb2 functions as a positive regulator of non-canonical Wnt signaling *a.o.* by repressing the expression of *Sfrp1*.

In mouse brain cortex interneurons (INs) comprise 20-30% of the neurons. They provide the crucial inhibitory neurotramsmitter GABA, which balances the exitatory glutamate output of the pyramidal neurons. Recently our lab showed, using first loss and thereafter gain-of-function

approaches, that Zeb2 also plays a role in interneuron (IN) guided migration (van den Berghe *et al.*, 2013). Zeb2 is present in the VZ and the mantle zones of the MGE, LGE, CGE and in POA and in the interneuron populations migrating along well-defined routes (Bartolini *et al.*, 2013; Métin *et al.*, 2006) to the cortex. Using Sip1|Nestin (Nestin-Cre mediated *Zeb2*, at that time still named *Sip1*, inactivation in the entire CNS; mice die at birth) and Sip1|Nkx2.1 (Nkx2.2-Cre mediated *Sip1* inactivation in POA and MGE with exception of the most dorsal part; mice are viable, but suffer from mycolonic seizures and/or motor deficiencies) the authors showed that Zeb2 is essential for the interneuron guided migration as the process is severely impaired in absence of Zeb2 and only very few INs reach the cortex.

Using focal electroporations in organotypic slice culture model van den Berghe and co-workers selectively inactivated *Zeb2* in small IN populations and demonstrated that neurons lacking Zeb2 (compared to the control slices/brain regions where Zeb2 was not deleted) displayed the migration defects, which indicated the cell-autonomous function of Zeb2 in the IN migration. Next, using RNA-seq the authors showed that the transcript levels of guidance cues (especially those of the Netrin/Unc5 system, in particular Unc5b) were deregulated in the *Zeb2* knockout brains. Furthermore, the authors confirmed (by RT-qPCR and ISH) that *Unc5b* (a short and long-range repulsive Netrin receptor, depending as to whether DCC joins in the receptor complex) was upregulated in the Sip1|Nkx2.1 model. They then demonstrated - by electroporating Unc5b shRNA expression construct in the organotypic Sip1|Nkx2.1 slice cultures with upregulated Unc5b - that the IN migration in the mutant was partially restored. Taken together Zeb2 was found essential for the IN guided migration of MGE-derived interneurons; also careful conclusions on subtle impact on cell fate of these mutant INs were found and discussed (van den Berghe *et al.*, 2013).

The second study (published back-to-back with the work of van den Berghe *et al.*, 2013) on the role of Zeb2 in the process of IN specification and migration – in collaboration with our lab – proposes Zeb2 more firmly as the crucial factor dictating the cortical versus striatal IN fate switch (McKinsey *et al.*, 2013). To understand Zeb2 function, the authors make use of two Cre-deletion mouse models: Nkx2.1-Cre (like in van den Berghe *et al.*, 2013), which deletes in the VZ of the MGE starting from E9.5, and Dlx1/2b-Cre, which drives Cre production in the SVZ and MZ of the entire subpallium from around E10.5. MGE is the source of both cortical and striatal interneurons during development. The homeobox TF Nkx2.1 plays the key role in specifying both cell types. The interneurons destined to become striatal maintain Nkx2.1 whereas those that migrate to the cortex loose Nkx2.1 expression. Upon *Zeb2* gene inactivation (using Nkx2.1-Cre), an increase in Nkx2.1 was observed and these INs failed to populate the cortex, migrated to the striatum and acquired a striatal-like fate instead. Similar results were obtained in the Dlx1;2b-Cre mouse model.

Next, the authors asked whether Dlx1 and Dlx2 - known to play an important role in the subpallial development including in interneuron migration to the cortex - could influence *Zeb2* expression. ISH of Zeb2 mRNA in the *Dlx1;2 null* mutants showed that *Zeb2* expression was strongly decreased in the subpallium (epecially in the SVZ), indicating that Zeb2 acts downstream of Dlx1/2, and that Dlx1/2 activate *Zeb2*. The authors also confirmed the direct binding of Dlx2 protein to candidate enhancer regions of Zeb2 by ChIP. Furthermore, microarray analysis of control, *Zeb2* KO (via Nkx2.1-Cre) and *Dlx1;2* conventional knockout samples, respectively, at E15.5 (i) revealed that *Dlx1;2* conventional KO and *Zeb2* conditional KO display similar IN-related gene expression changes as compared to the control samples, (ii) proposed *Maf*, *MafB* and *Cxcr7* mRNA as new signature for migrating interneurons and (iii) showed that upon Zeb2 inactivation (or *Dlx1;2* deletion) the expression of these markers is lost hence the IN cell identity is perturbed. Taken together, these two studies confirmed the crucial role of Zeb2 in IN guided migration and cell specification, while the second study demonstrated that Zeb2 acts directly downstream of Dlx1/2 and that Dlx1/2-Zeb2 axis is crucial for migrating IN specification.

1.4.4.2 Hematopoiesis

Hematopoietic stem/progenitor cells (HSCs/HPCs) give rise to many different cell types present in the adult blood. Hematopoietic development in the mouse occurs in two waves, i.e. the primitive and the definitive wave (Orkin and Zon, 2008). The primitive wave takes place in the yolk sac at around E7.5. A second induction takes place in the dorsal aorta and is followed by hematopoiesis in the dorsal aorta – gonadal – mesonephros (AGM) region with the emergence of the first cells that will become definitive HSCs (Medvinsky and Dzierzak, 1996; Müller *et al.*, 1994). Next, around E12, HSCs migrate to the fetal liver, and this migration wave is followed by the migration to the bone marrow (around birth) that later becomes the source of HSCs throughout life (reviewed in Mikkola and Orkin, 2006).

Zeb2 is also present in adult HSCs (self-renewing) and hematopoietic progenitor cells (HPCs, non-self-renewing multipotent progenitors) and in aortic hematopoietic clusters from the earliest stages of definitive embryonic hematopoiesis. To assess the role of Zeb2 in embryonic hematopoiesis, our lab – in collaboration with J. Haigh (Gent) and E. Dzierzak (Rotterdam) – knocked-out *Zeb2* using the Tie2-Cre *deleter* mouse (deletes in all endothelial cells and their progeny including HSCs, and in all definitive blood cells). No obvious defects were present before E11.5 suggesting that Zeb2 is not essential for primitive hematopoiesis. From E11.5-E12.5 the mutant embryos showed severe hemorrhages in the brain and heart malformations. Loss of Zeb2 (Zeb2 $^{-/Tie2\text{-}Cre}$) resulted in embryonic lethality around E12.5-13.5.

To address the role of Zeb2 in adult hematopoiesis, the authors inactivated *Zeb2* in hematopoietic stem cells after the initial stem cell formation, using the Vav-iCre mouse line. Most of the resulting $Zeb2^{-/Vav\text{-}iCre}$ mice died at birth and displayed intracranial bleedings resembling the $Zeb2^{-/Tie2\text{-}Cre}$ phenotype. CFU assays on HSCs derived from $Zeb2^{-/Tie2\text{-}Cre}$ E10.5 yolk sacs, E11.5 AGM regions, E11.5-E12.0 fetal livers, and E12.0 peripheral blood were analyzed and all showed a dramatic impairment in hematopoietic colony formation compared with the controls, indicating a severe block in hematopoietic differentiation in all lineages. Also, a similar impairment of hematopoiesis in the fetal livers of $Zeb2^{-/Vav\text{-}iCre}$ mice was observed. Consistent with these findings Zeb2-deficient ESCs were unable to undergo hematopoietic differentiation *in vitro*.

The authors also showed that loss of Zeb2 alters HPCs located in the fetal liver probably due to increased adhesive properties (accompanied by an increase in Integrinβ1, Cxcr4 and E-cadherin levels), leading to a defective homing and hence a decrease in the HSC numbers within the bone marrow. Lastly, the authors showed that hematopoietic-specific loss of Zeb2 leads to a decrease in Ang1/Tie2 and affects pericyte coverage of the vasculature in the developing brain. This could explain the severe hemorrhages in the *Zeb2* mutant embryos studied here. Taken together, this study addressed for the first time role of Zeb2 in hematopoiesis and showed that Zeb2 is dispensable for the formation of HSCs/HPCs, but is essential for their differentiation potential and stemness (Goossens *et al.*, 2011). This work, addressing the role of Zeb2 further in adult hematopoiesis, is now continued in collaboration with J. Haigh (Melbourne) and V. Janzen (Bonn).

1.4.4.3 Melanocyte development

Melanocytes are specialized and pigmented skin cells that produce melanin responsible for skin and hair color and for protection against UV light. During development melanocytes originate from the neural crest cells. These are transient multi-potent cells that arise by delamination from the developing dorsal neural tube between E8.5 and E10.5 (via an accompanying EMT process) and later form the immature glial cells innervating the skin. Mitf (Microphthalmia-associated transcription factor) is the key regulator of melanocyte development and induces melanocyte differentiation and initiates pigment production (reviewed in Ernfors, 2010).

Zeb2 is present in the melanocytes of skin, as well as in the differentiated melanocytes of hair follicles. To address the role of Zeb2 during melanocyte development our lab – in collaboration with G. Berx (Gent) - deleted it in the melanocyte lineage using the Tyr hydroxylase-Cre mouse (deletes from

E10.5). *Zeb2* inactivation leads to loss of hair pigmentation and severely impaired melanoblast development and differentiation. Moreover, Mitf was severely reduced in absence of Zeb2, indicating that Zeb2 is an important regulator of Mitf-dependent melanocyte differentiation. Zeb2 KD (by siRNA) in cultured immortalized melanocyte cells (Melan-a line as well as the human SKMel28 melanoma cell line) leads to a similar phenotype. Mitf overexpression in the Zeb2 KD lines rescued the melanocyte differentiation defect, suggesting that the process is Mitf-dependent. Interestingly, Zeb2 down regulation that leads to Mitf down regulation and impairs melanocyte differentiation, was inversely correlated with Zeb1 levels, which were found increased upon *Zeb2* KO (or Zeb2 KD). Zeb1 is present in melanocyte stem cells and an increase in Zeb1 level may indicate that the Zeb2-depleted cells maintain their stemness. Taken together this data shows that Zeb2 is involved in the regulation of the levels of Mitf mRNA and thereby coordinates the development and differentiation of the melanocyte lineage.

1.5 Zeb2 in disease

1.5.1 Mowat-Wilson syndrome

Heterozygous mutations in *Zeb2* cause Mowat-Wilson syndrome (MOWS; OMIM ♯235730). MOWS was recognized for the first time in 1998 (Mowat *et al.*, 1998) as an independent syndrome and the genetic link was described by Cacheux and co-workers (Cacheux *et al.*, 2001) and by the Wakamatsu team (Wakamatsu *et al.*, 2001). All MOWS patients display moderate-to-severe intellectual disability and delayed motoric development. They often present a wide range of additional developmental anomalies in eye and renal tract development (>50%), agenesis of corpus callosum (>50%), congenital heart defects (>50%), epilepsy (70%) (Cordelli *et al.*, 2013) and Hirschsprung disease (>50%) (reviewed in Zweier *et al.*, 2005). The patients are diagnosed based on typical facial features combined with intellectual disability. The facial features of MOWS patients include hypertelorism, thick eyebrows with lateral flaring, open mouth expression and posteriorly rotated ear lobes often described as “donut-shaped” (see Fig. 11). Currently, there are more than 200 patients diagnosed in detail in the world and in most of them the *Zeb2* mutation has been mapped. To date, over 100 different mutations have been described (Garavelli and Mainardi, 2007). Mutations are scattered around the gene with a large proportion in exon 8, which is the biggest exon of Zeb2 (and the equivalent of exon 7 used for targeting in our KO mouse models) and there is no clear genotype-phenotype correlation (Dastot-Le Moal *et al.*, 2007; Zweier *et al.*, 2005).

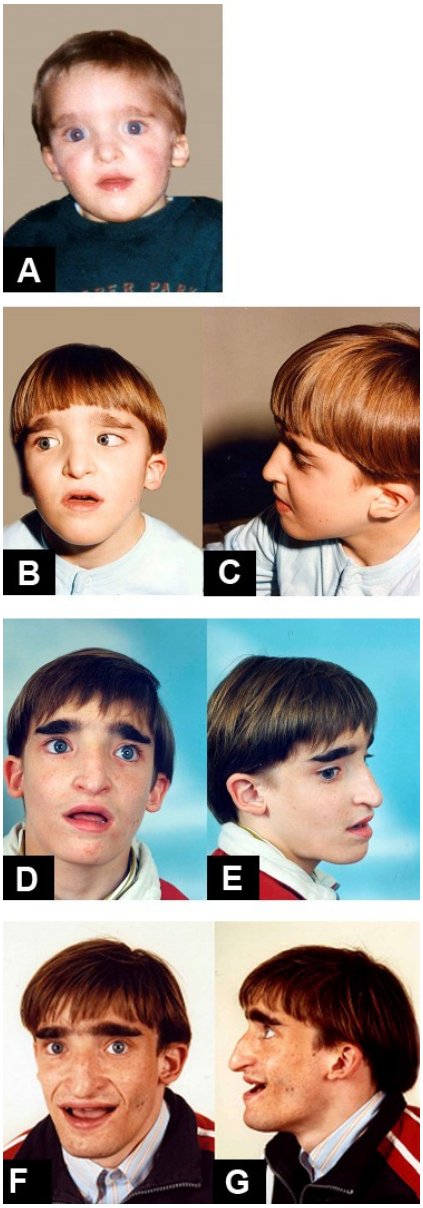

Figure 11. Mowat-Wilson syndrome, highlighting some of the clinical features of MOWS patients evolving with age.

The ages shown are: (A) 1 year and 6 months; (B-C) 5 years; (D-E) 13 years and 8 months; (F-G) 18 years.(Garavelli and Mainardi, 2007)

1.5.2 Cancer

Epithelial-to-mesenchymal transition (EMT) is a key cellular event in cancer spreading and it is a major contributor of aggressive cancer metastasis originating from epithelial-derived cancers, which is the most prevalent class. During EMT epithelial cells lose their cell-cell adhesions and apical-basal polarity and acquire mesenchymal phenotype what allows them to invade the extracellular matrix.

One of the key features of EMT is the downregulation of epithelial cadherin (E-cadherin, E-cad, Cdh1) which is a tumor invasion suppressor (Vleminckx *et al.*, 1991). E-cadherin gene expression is regulated by a number of known transcription factors (TFs), including Twist, Snail family members, and Zeb1 and Zeb2. Zeb2 was shown to downregulate *E-cadherin* and contribute to loss of cell adhesion and hence more invasive phenotype in a number of cell lines (Comijn *et al.*, 2001). Moreover, a number of studies have documented elevated Zeb2 mRNA and/or protein levels in specific types of cancers, negatively correlating with E-cad levels and indicating bad prognosis, including due to increased invasion, with high Zeb levels. These cancers include primary colorectal cancer and liver metastasis (Kahlert *et al.*, 2011), gastric cancer (Rosivatz *et al.*, 2015), hepatocellular carcinoma cells (Miyoshi *et al.*, 2004) and ovarian and breast carcinomas (Elloul *et al.*, 2005).

Recently, Zeb2 was shown to be involved in a reciprocal negative feedback regulation together with the miR-200 cluster (reviewed in Brabletz and Brabletz, 2010; Bracken *et al.*, 2008). miR-200 down tunes Zeb2 (and the invasive phenotype of $Zeb2^{high}$ cells) in the NCI-60 cancer cell line (Park *et al.*, 2008). It was also shown that NMuMG cells avoid EMT by enhancing E-cadherin levels through direct targeting of Zeb2 by miR-200 (Korpal *et al.*, 2008). Taken together, this data strongly suggests that (i) the balance between Zeb2 and miR-200 members co-determines the epithelial versus mesenchymal phenotype and hence aggressiveness of tumor cells, and (ii) that Zeb2 is involved in cancer progression and metastasis. This interaction with *a.o.* miR-200 may also be relevant for Zeb2's functions in ESCs studied here.

2 Chapter 2: Objectives

The overall objective of my PhD research was to contribute to a number of stem cell-based and *in vivo* studies (in the mouse) aiming at understanding (i) what Zeb2's role is in selected processes in embryogenesis (*functional analysis*) and (ii) ideally in such context, combine this with the question how Zeb2 regulates cell fate decisions during development (*action mechanism(s)*). Therefore, the focus of the first part of my PhD thesis is on the role of Zeb2 in mouse brain development performed in collaboration with many team members, and the second part – and more extensive, and predominantly own work (together with the bio-informatics done by R. Dries) - will do so for Zeb2's role in cultured mouse ESCs, in particular in their transition from pluripotent state cells to commitment and subsequent differentiation into mesoderm, endoderm and neuroectoderm lineages. The respective specific aims of my research can be summarized as follows:

2.1 To identify and validate Zeb2-dependent and/or candidate target genes during mouse brain cortex development

In the first part of my PhD I joined an ongoing pioneer project within the laboratory that took us into functional studies in CNS development, more specifically the embryonic forebrain, for the very first time. This work ultimately led to unraveling the molecular mechanism by which Zeb2, surprisingly – also for the cortex development field - as a non-autonomous factor, regulates the timing of neurogenesis and gliogenesis in brain cortex development (Seuntjens *et al.*, 2009). My specific contribution was to identify Zeb2-dependent and/or candidate target genes candidate during cortex development and amongst these identify the genes that help to mechanistically explain the cortical phenotype observed in a (growing) series of conditional *Zeb2* KO mice and validate these candidates subsequently by a combination of *a.o.* ChIP, RT-qPCR and ISH (see Chapter 3, sections 3.1-3.4).

2.2 To develop an embryonically relevant cell culture system to assess Zeb2's function and action mechanism

To gain more insight into the likely multiple mechanisms of action of Zeb2 during neurogenesis, my aim was to establish a simplified, robust cell differentiation system. My choice was to use mouse ESCs and – based on a number of *in vivo* observations in the lab's parallel projects in mouse embryos – model neurogenesis *in vitro.*

The ESC system I used throughout (for a number of reasons) is based on embryoid body formation and retinoic acid-based neural induction; it yielded pure neural cell populations in control ESCs in a period of 15 days. Griet Verstappen, Leo van Grunsven and I next generated *Zeb2* KO ESCs and I subsequently subjected these (together with control ESCs) to the neural differentiation protocol. I observed that *neural development* in absence of Zeb2 was abolished.

To validate in parallel whether absence of Zeb2 leads also to defects in differentiation towards mesodermal and endodermal lineages, we subjected *Zeb2* KO and control lines also to an acknowledged *general* EB-based differentiation protocol. We observed that differentiation to all three lineages was impaired indicating that Zeb2 plays a role in early cell fate decisions, possibly including exit from pluripotency in ESCs (see Chapter 4, sections 4.1-4.3).

2.3 To identify Zeb2-dependent genes during *in vitro* neural differentiation using transcriptomics

As an important part of my PhD project, made possible via the establishment of the aforementioned ESC system, I applied RNA-sequencing to identify Zeb2-dependent genes at three essential time-points: d0 (pluripotent stem cell, very low Zeb2 levels), d4 (multipotent progenitor, low Zeb2 levels) and d6 (early neural progenitor, high Zeb2 levels). Gene ontology analysis confirmed that neural development in the *Zeb2* KO ESCs was affected, but also that expression of genes meanwhile linked to pluripotency and DNA-methylation was deregulated in those cells on d0, d4 and d6. As this part of my work was performed in a period where major new insights and omics approaches were used in the field, I continuously applied new types of analysis in order to meet the increasingly stringent requirements in the field (see Chapter 4, sections 4.3.1-4.3.9)

2.4 To compare dynamic changes in DNA-methylation during Ctrl and *Zeb2* KO ESC differentiation

Our RNA-seq data revealed that many of the DNA-methylation and hydroxylation genes were deregulated in *Zeb2* KO. We asked whether this transcriptional deregulation had an effect on the actual DNA-methylation status of the *Zeb2* KO cells in pluripotency and during differentiation. Temporal reduced representation bisulfate sequencing revealed that *Zeb2* KO cells initially correctly acquire methyl marks but do not maintain those later during differentiation what is in line with our observations on transcriptional changes in genes related to DNA-methylation and pluripotency in the *Zeb2* KO cells (see Chapter 4, sections 4.3.5-4.3.6).

2.5 To validate selected candidate direct target genes and gain insight into possible mechanisms of action of Zeb2 during ESC differentiation

My comparative RNA-sequencing data yielded large numbers of deregulated genes in *Zeb2* KO ESCs at each of the three time points analyzed. I decided to then focus on maintenance of and exit from pluripotency and on methylome-related genes since our working hypothesis rapidly took us to the defects in pluripotency exit as primary cause for the subsequent failure in cell lineage commitment. Doing so, we also hoped to contribute to new insight into pluripotency exit in particular (see Chapter 4, sections 4.3.4, 4.3.7).

Taken together, I managed to show that *Zeb2* KO ESCs do not downregulate acknowledged pluripotency genes during differentiation and that these cells retain their self-renewal capacity over prolonged periods of time. In addition, I was able to show show that *Nanog* and *Cdh1* are direct targets of Zeb2 in ESCs. I also identified Tet1 as a Zeb2-dependent gene and demonstrated that, upon Tet1 knockdown, the Zeb2-deficient ESCs can again exit from pluripotency and partially regain their differentiation capacity.

I also identified REST (a neural-inhibitory gene) as another gene whose expression depends on Zeb2. After investigating this further, I propose a second action of Zeb2 through which it may regulate neurogenesis (see Chapter 4, section 4.3.8).

2.6 To explore the neurodevelopmental potential of Zeb2 domain mutant and Zeb1 knock-in mouse ESC lines

To understand how Zeb2 functions during neural development and which of the many new Zeb2 protein partners (identified in parallel work in our laboratory) are crucial for its multiple roles during pluripotency exit and in neurogenesis we generated – in collaboration with the team of Jody Haigh (Gent, then Melbourne) and Steven Goossens (Gent) – by homologous recombination a series of Zeb2 domain mutant ESCs that express the transgene Zeb2 (wild-type, see above, and mutant) cDNA from the *R26* locus (*i.e.* under the control of the endogenous R26 promoter) in a *Zeb2* KO ESC context.

Importantly, the insertion of wild-type (WT) Zeb2 cDNA (both mouse and human) were able to rescue the neural differentiation defect seen in *Zeb2* KO cells. Interestingly, ESCs with the insertion of the mutant *Zeb2* (with a deletion in its Smad-binding domain, SBD) performed better in the neural differentiation protocol yielding more pure neural cell populations than the *R26*-driven WT Zeb2. This indicates also that Zeb2-Smad co-operation (possibly via direct interaction) has an inhibitory effect on neurogenesis. In line with this, insertion of wild-type Zeb1 (which does not detectably bind to Smads in our hands) in the *R26* locus of the *Zeb2* KO ESCs also efficiently rescues the neurodevelopmental potential of the *Zeb2* KO cells (see Chapter 4, section 4.3.12). To obtain a full picture of how Zeb2 functions in concert with its (meanwhile more, Conidi *et al.*, *unpublished results*) partners, this initial work in this PhD research provides the possibility of generating additional mutant lines with deletions in *e.g.*. the NuRD-interaction motif and other motifs, including protein partner binding ones, in Zeb2.

3 Chapter 3: Role(s) of Sip1/Zeb2 in the development of the mouse brain cortex

Parts of this Chapter were published in: *Nat Neurosci*. 2009 Nov;12(11):1373-80. doi: 10.1038/nn.2409.

In the original publication, the nomenclature used was Sip1. We kept this initial nomenclature in the figures and legends in Chapter 3 in order not to modify the figures from the published paper, whereas in the text, for keeping the consistency throughout the thesis, we use Zeb2. Chapter and section titles however use double "Sip1/Zeb2" nomenclature.

3.1 Introduction

During brain cortex development its pyramidal neurons are born in an orderly sequence, which ultimately – after radial migration of the neuronal cells – is reflected by the ordered, multi-layer nature of the mammalian brain cortex (see Chapter 1). Cell proliferation and specification is orchestrated here by both intrinsic and extrinsic signals.

The first studies addressing cortical cell specification during corticogenesis demonstrated that cell fate is encoded within the individual neural progenitor cells (NPCs). Experiments where such progenitor cells from early embryonic brains were transplanted to developmentally older brains (and *vice versa*) showed that their developmental potential decreases with gestation time and that the fate of these progenitors is to a large extend, and sometimes claimed exclusively, intrinsically determined. The early progenitors transplanted in the older brains could produce late-born neurons, however the late progenitors had lost their potency with time and could only generate late-born neurons, regardless of the environmental cues provided by the host. In addition, it was shown in parallel experiments that cortical pyramidal neuron progenitors retained their neurogenic timing *in vitro* and that the developmental potency of even co-cultured cells from different cortical stages is not dictated by the cell environment (Desai and McConnell, 2000; Frantz and McConnell, 1996; McConnell and Kaznowski, 1991; Shen *et al.*, 2006). Interestingly, the developmental potential is restricted rather gradually: the progenitor cells that give rise to neurons of a particular cortical layer maintain their competence for longer than needed for the production of that layer only, suggesting that an additional cell non-

autonomous mechanism plays a role in the coordination of the appropriate timing of neurogenesis (Desai and McConnell, 2000).

Regulation of development by signaling proteins that act on cells directly to induce cellular responses in a concentration-dependent manner has long been known as one of the crucial mechanisms that determine cell fate and steer embryonic tissue patterning as well (reviewed in Ashe and Briscoe, 2006), but the concept of extrinsic regulation of corticogenesis was introduced rather recently. The extrinsic queues in the brain originate locally, *e.g.* from within the progenitor zone, but could during development also come from remote sources such as blood vessels, meninges or cerebrospinal fluid in the neighborhood of the VZ.

Retinoic acid (RA) produced by the meninges was shown to regulate neuron generation in the cortex. The transcription factor Foxc1 is – in addition to other tissues – present in the meninges and head mesenchyme. Genetic removal of *Foxc1* leads to severe defects in meninges development, which lead to abnormalities in skull and cerebral cortex formation with very prominent cortical layer disorganization, reduction in intermediate progenitors and neuron number, as well as in overall elongation of the neuroepithelium (Zarbalis *et al.*, 2007). In absence of Foxc1, the mRNA expression of a number of genes involved in RA-signaling, including *Rdh10* and *Raldh2* that are crucial for atRA (all-*trans* retinoic acid; Sandell *et al.*, 2007) synthesis, is also reduced. Interestingly, injection of atRA into pregnant mice carrying *Foxc1* mutant embryos partially rescues the *Foxc1* phenotype. Treatment with atRA leads to a decrease in the length of the neuroepithelium and increase in numbers of early and late-generated neurons. The cortex still showed defects in the laminar organization that are most probably caused by the meningeal/basement membrane defects that are present in atRA-treated mutants. This important study shows that corticogenesis not only depends on the cell-intrinsic program but is dictated also by extrinsic input (here atRA) from the surrounding meningeal cells (Siegenthaler *et al.*, 2009).

Pals (Protein associated with Lin7) and Pten (Phosphatase and tensin homolog), both present in the progenitor zone of the cortex, play opposing roles in cortex development and balance the responses to polypeptide growth factors present in the CSF. Genetic inactivation of *Pals* leads to a decrease in progenitor cell numbers and premature cell-cycle exit, coupled to excessive production of deep layer neurons followed by massive cell death, which collectively induces abrogation of the entire cortex (Kim *et al.*, 2010). Knockout of *Pten* results in enlarged brain size due to increased cell proliferation and decreased cell death (Groszer *et al.*, 2001). Interestingly, Pten deficiency in the *Pals1* mutant background results in an almost normally sized brain. Both proteins have been shown to regulate the localization of IgfR to the apical surface of the cells lining the ventricles of the developing forebrain.

Manipulation of Igf/IgfR signaling results in phenotypes similar to those observed in Pals/Pten defeciencies: IgfR deficiency causes a decrease in the progenitor proliferation and microcephaly (Liu *et al.*, 2009), while induction of Igf signaling results in brain overgrowth and macrocephaly (Ye *et al.*, 2004). IgfR responds to CSF-containing Igf2, which promotes progenitor proliferation. These studies show how Pals and Pten spatially restrict and expand, respectively, *IgfR* expression in the progenitor zone to modulate the cellular response to Igf2 signals present in the CSF (Lehtinen *et al.*, 2011). Various other trophic factors have meanwhile been shown to modulate cell proliferation and migration, but for many of these the mechanism(s) of action remain to be clarified. Bmp7 present in the choroid plexus, meninges and the cortical hem is one of such factors that is present in the CSF. Bmp7 was found to regulate fundamental properties of radial glia and neural progenitors and their survival. In a collaboration with the team of Daniel Graf (in Zürich, presently in Alberta) our laboratory could show that in the absence of Bmp7 the overall thickness of the cortex was reduced and that the radial glia no longer attached to the meninges (Segklia *et al.*, 2012). It was postulated that Bmp7 exerts its function via regulating the well-characterized neurogenic Pax6/Ngn2 axis in the developing cortex (Scardigli *et al.*, 2003; Segklia *et al.*, 2012).

Brain-derived neurotrophic factor (BDNF, a neurotrophin) is one of the key extrinsic factors that together regulate laminar formation in the developing brain cortex. Administration of BDNF into the VZ of E13.5 mouse brains shortens the cell cycle of the progenitors in the VZ by accelerating completion of the S-phase of the cell cycle what causes a more rapid release to the basal VZ and their migration. This results in a switch from the generation of layer 4 neurons to neurons typical for deep layers 5 and 6, whereas administration of a function-blocking anti-BDNF antibody induces a fate switch and the neurons acquire fate of the upper layers 2/3 (Fukumitsu *et al.*, 2006).

3.1.1 Sip1/Zeb2 levels are high in post-mitotic neocortical cells

Zeb2 transcripts (not shown) and protein (Fig. 12) are detected during neurogenesis in the post-mitotic cells of the mouse cortex, with barely detectable levels in progenitors. (Fig.12a–c). In the medial/cingulate cortex, the Zeb2 domain at late gestation (E18.5) became restricted to cells of the deep layers (Fig. 12c).

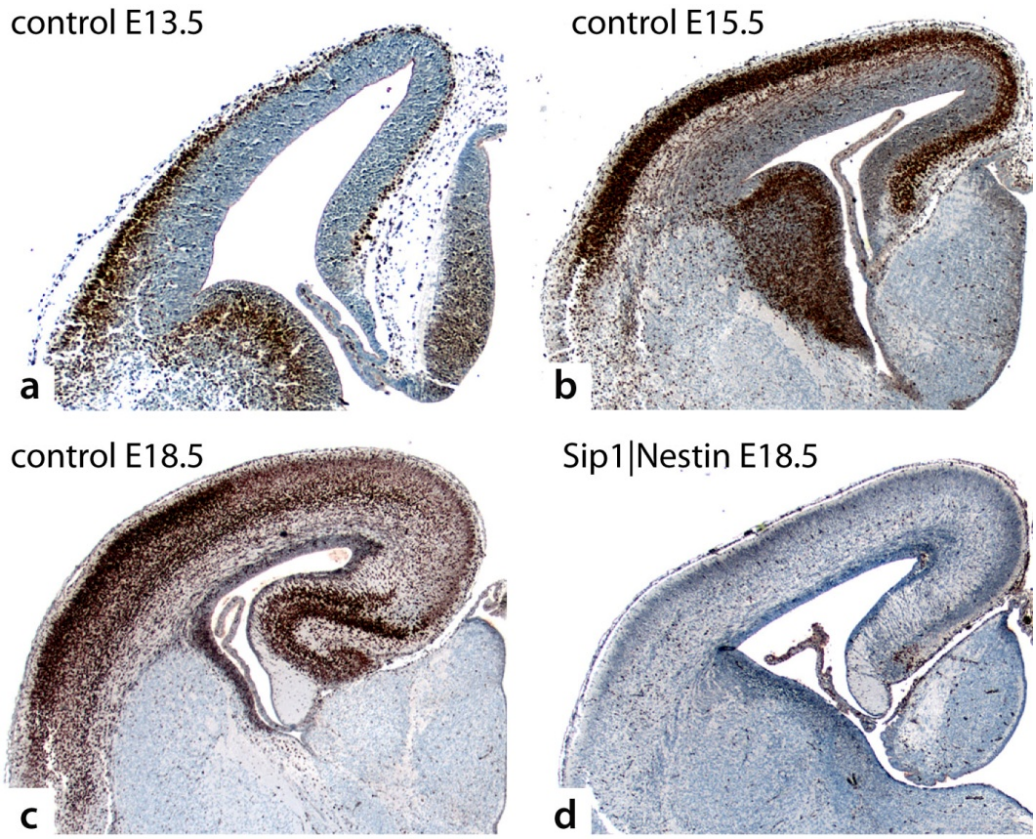

Figure 12. Expression of Sip1/Zeb2 in the embryonic cortex.

(a) Sip1 immunostaining at E13.5 was found in the ventricular zone of the ventral telencephalon and in the post-mitotic cells of the cortex. (b) At E15.5, Sip1 expression was intense in the post-mitotic cells of the cortex, where invading Sip1-positive cells from the ventral telencephalon could be seen as well. (c) At E18.5, Sip1 synthesis was area specific; in the lateral cortex, it was maintained in the entire cortical plate, whereas it was downregulated in the upper layers of the cingulate cortex. (d) Conditional knockout of Zeb2 (shown here using Nestin-Cre) resulted in the almost complete loss of Zeb2 protein from the cortex and the ventral telencephalon. (Seuntjens *et al.*, 2009)

3.1.2 *Sip1/Zeb2* knockout mouse models

To document the effects of Zeb2 removal on cortex development, three Cre-*deleter* strains were used: Nestin-Cre to inactivate Zeb2 in the entire CNS, Emx1-IRES-Cre to ablate Zeb2 exclusively in neocortical progenitors starting from E9.5, and Nex-Cre to separate Zeb2's function in early cortical progenitor cells from its role in post-mitotic neurons. Zeb2|Nestin mice died at birth, whereas both Zeb2|Emx1 and Zeb2|Nex mice survived until 3-4 weeks after birth. Notably, the neocortical phenotype observed in all three models was nearly identical.

3.1.3 Lack of Sip1/Zeb2 causes premature generation of upper layers at the expense of deep layers

The absence of Zeb2 leads to a reduction in numbers of deep layer neurons (See Fig. 13) marked by Ctip2 (layer 5) and Tbr1 (layer 6) (Fig. 13 a-c), whereas upper layers marked by Satb2 (layers 2-4, Fig. 13 d-f) are significantly expanded at E17.5 (Fig.13h).

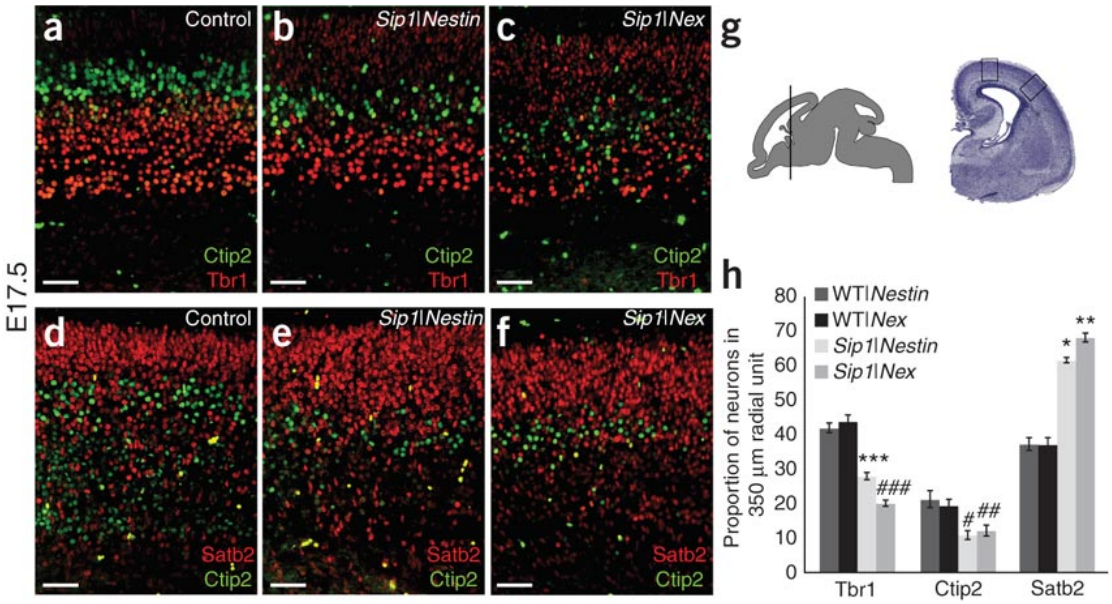

Figure 13. Sip1/Zeb2-deprived brain cortex contains excessive numbers of upper layer neurons at the expense of deep layer neurons whilst maintaining their relative position within the cortex.

(a–c) Double immunolabeling for Tbr1 (layer6) and Ctip2 (layer5) showed a reduction of deep layers when Sip1 was removed from either the entire brain (Sip1|Nestin; b) or post-mitotic neurons (Sip1|Nex; c). (d–f) Double labeling of Satb2 (layers 2–4) and Ctip2 showed increased size of upper layers. (g) The regions (boxes in right panel) chosen for quantification in sagittal and coronal sections of mouse E17.5 brain are indicated, as well as the plane of sectioning used for quantification. (h) Quantification of the proportion of neurons of a certain layer in the total number of cells present in the cortical plate. # P = 0.006, ## P = 0.00057, ### P = 0.000051, * P = 0.00014, ** P = 0.000057, *** P = 0.000063. Error bars indicate s.e.m. Scale bars represent 50 µm. (Seuntjens *et al.*, 2009)

A birth-dating study done by immunolabeling mitotically active precursors with BrdU (by intraperitoneal injections of the pregnant females) in combination with layer-specific markers showed that the peaks of production of Ctip2+ layer 5 neurons and Brn2+ layer 2–4 cells had shifted to 1 day (d) earlier in *Zeb2* mutants (Fig. 14 a,b,c), whereas the generation of Tbr1+ layer 6 neurons remained unchanged (Fig. 14 a). This suggests that the onset of neurogenesis in *Zeb2* conditional mutants was not altered, but subsequent neuronal subtypes were prematurely generated, indicative of a forward shift of neurogenesis. This is confirmed by staining for Brn2, especially at E14.5 (Fig. 14).

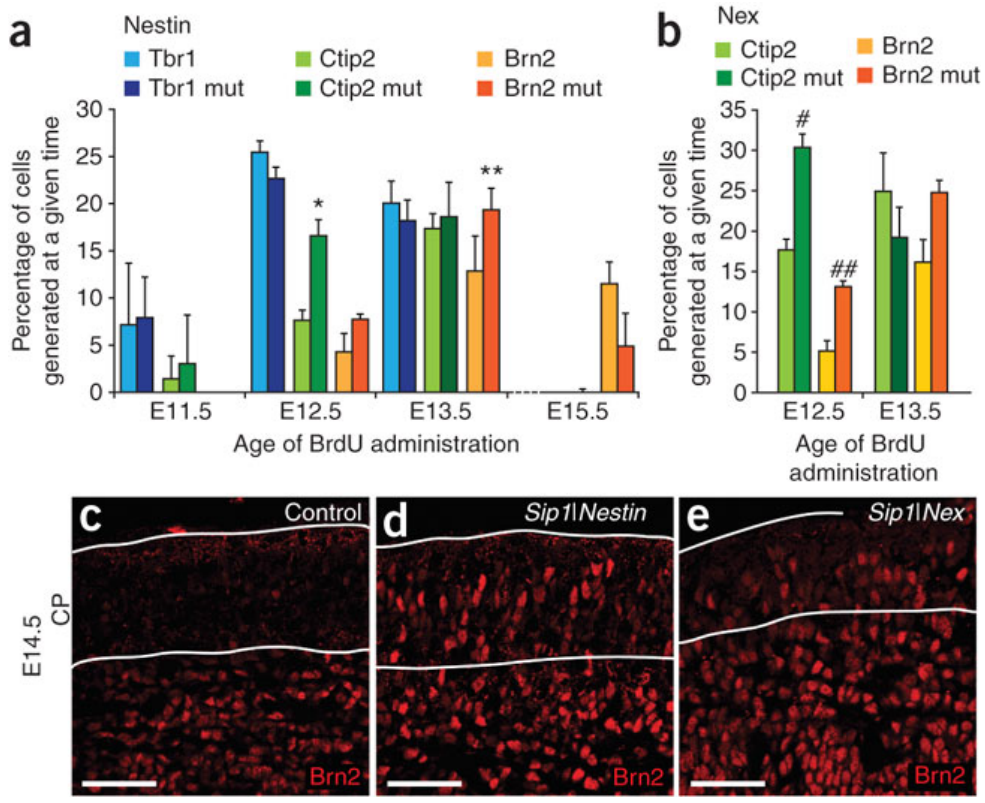

Figure 14. Sip1/Zeb2 deletion in neocortical post-mitotic cells causes premature generation of layer 2-5 neurons.

(a) Pulse-chase experiments with BrdU administered at different time points during development (shown on the x axis) followed by analysis of the indicated layer-specific markers at E18.5. Quantification of percentages of cells that were stained for both BrdU and the layer-specific marker at the time of analysis, corresponding to different time points of BrdU administration in control and Zeb2|Nestin (mut) mice. There was no difference in the timing of production of Tbr1-positive neurons. A higher proportion of Ctip2-expressing neurons incorporated BrdU label at E12.5 (* P = 0.002) and more Brn2-positive cells were generated at E13.5 (** P = 0.036) in the mutant cortex. At E15.5, no layer 5 and 6 neurons were born in either the control or mutant, whereas the generation of layer 2–4 neurons was nearly complete in the latter. Error bars indicate s.e.m. (b) Similar pulse-chase experiments at E12.5 and E13.5 in Zeb2|Nex and control mice showed increased generation of layer 2–5 neurons (# P = 0.0047 and *## P = 0.013) at E12.5. Error bars indicate s.e.m. (c–e) Immunostaining for upper-layer marker Brn2 at E14.5 revealed neurons with upper-layer characteristics in the mutant (Zeb2|Nestin and Zeb2|Nex) cortical plate at a stage when such neurons were absent in the control. Scale bars represent 50 µm. (Seuntjens *et al.*, 2009)

3.1.4 Sip1/Zeb2 is also crucial for appropriate control of the timing of gliogenesis

Gliogenesis in the mouse cortex starts after the onset of neurogenesis, *i.e.* at E17.5, and proceeds beyond birth. We observed that in the *Zeb2* KO brain there were significantly more GFAP+ cells dcompared to control (Fig. 15 e,f). Staining with anti-Ki67 (which marks proliferating cells) showed that the number of dividing cells at this stage was significantly higher in the *Zeb2* mutant brains and that more dividing cells were present in the IZ (and not in the VZ) of the *Zeb2* KO brains as compared to the control at E17.5 (Fig. 15 a-d). BrdU pulse-chase experiments done to document further the precocious generation of glial cells in the Zeb2-deficient brains revealed that at E16.5 about 45% more Olig2+ cells

were generated in mutant compared to control embryos (Fig. 15 h,j,l,n). In addition, we detected more cells born at E16.5 that differentiated into GFAP+ astrocytes at P2 and P4 in the mutant cortex (Fig. 15, g,i,k,m). In conclusion, the shift forward in neurogenesis in *Zeb2* KO mice paves also the way for precocious and enhanced proliferation of glial precursors, which in turn leads to the production of higher numbers of astrocytes at early postnatal stages.

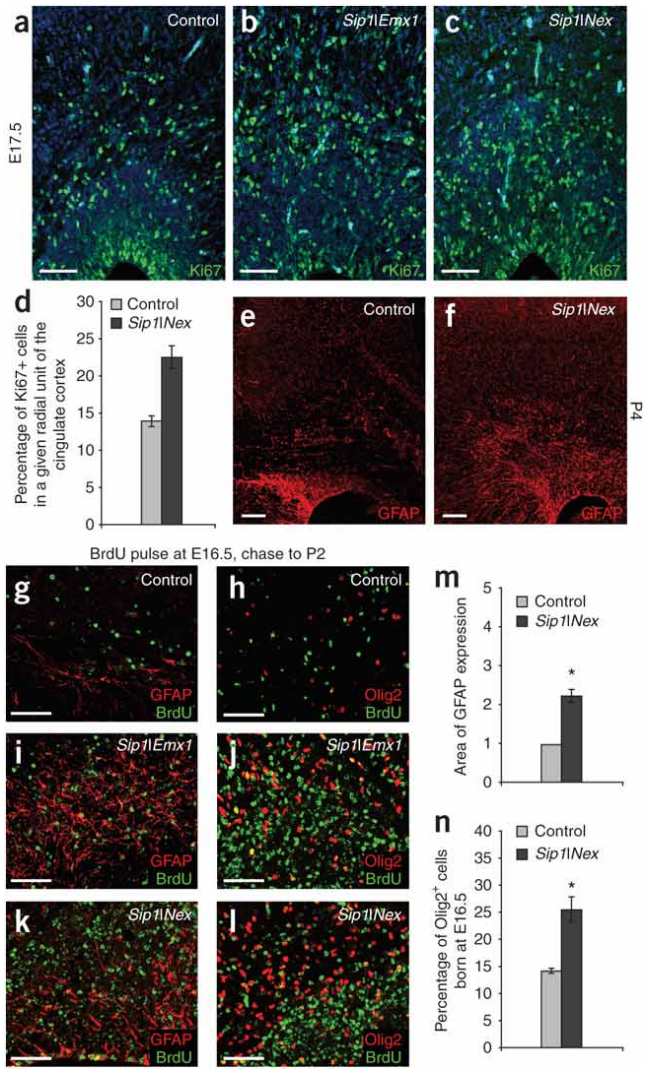

Figure 15. Enhanced astrocytic proliferation and premature and increased gliogenesis in Sip1/Zeb2 conditional mutants.

(a–c) Shown here is the expression of proliferation marker Ki67 at E17.5 in control (a), Zeb2|Emx1 (b) and Zeb2|Nex (c) mice in a region close to the medial cortex. (d) The increase in proliferation in the mutant cortex was quantified over an entire radial unit and is represented as the percentage of the total number of cells counted in this unit that were labeled for Ki67. * P = 0.0063. (e,f) Expression of astrocytic marker GFAP was found to be twofold higher in the cingulate cortex in Zeb2|Nex at P2 (immunostaining for GFAP at P4 shown in e and f; quantification in m). (g–l) To trace the origin of these astrocytes, we pulse-labeled dividing cells in the cortex with BrdU at E16.5 and chased them at P2. Double immunostaining for BrdU and GFAP (g,i,k) and for BrdU and Olig2 (h,j,l) indicated that precocious production of glial cells was occurring in Zeb2|Emx1(i,j) and Zeb2|Nex (k,l) conditional knockout mice. Scale bars are 50 µm. (m) We measured the area of GFAP expression in a specific region of the cortex in both control and mutant and normalized the mutant values with regard to the control. * P = 0.0018. Error bars indicate s.e.m. (n) Twice as many Olig2-expressing glial progenitors were born at E16.5 in Zeb2|Nex brains as in control. * P = 0.0077. Error bars indicate s.e.m. (Seuntjens *et al.*, 2009)

3.2 Methodology and Materials

3.2.1 Expression profiling by microarray analysis

In the first series of arrays, we examined gene expression in cortex and hippocampal tissue from two control (Zeb2$^{loxP/+}$; *Nestin*-Cre) and two mutant (Zeb2$^{loxP/-}$; *Nestin*-Cre) E18.5 littermates, each taken from two different litters. In a separate experiment, we used E14.5 cortex and hippocampal tissue from two control (Zeb2$^{loxP/+}$; *Nex*-Cre) and two mutant (Zeb2$^{loxP/-}$; *Nex*-Cre) littermates. RNA from cortex and hippocampus tissues was isolated using an RNeasy kit according to the manufacturer's protocol (Qiagen). The isolated RNA was inspected for integrity and purity using an Agilent Bioanalyzer and a NanoDrop spectrophotometer, respectively. All samples were of similar RNA quality. Starting with 1 µg of total RNA, we amplified RNA by *in vitro* transcription with a biotin labeling reaction (Amersham Biosciences). The probes were purified and analyzed again for yield (>20 µg) and purity (260:280 and 260:230 nm). We fragmented 10 µg of the resulting antisense RNA according to the manufacturer's protocols (Amersham Biosciences) and resuspended it in 260 µl of hybridization buffer.

Codelink Mouse Whole Genome array is a single array representing ~35,000 transcripts. The gene array ChIPs were hybridized in a shaker-incubator at 37 °C at 300 rpm for 18 h and washed and stained with Cy5-streptavidin according to manufacturer's protocols (Amersham Biosciences). The Agilent DNA Microarray scanner was used for scanning and image analysis was performed with Codelink Expression Analysis 4.1 software (Applied Microarrays).

3.2.2 RT-qPCR and ISH

The expression of *Nt3* and *Fgf9* mRNA was validated by RT-qPCR (ABI Prism 7000, Applied Biosystems) on E14.5 cDNA synthesized from mutant and control cortex mRNA. *Nt3*, *Fgf9*, *Fgfr2*, *Fgfr3* and *TrkC* mRNA was detected on brain slices (6 µm) by cold ISH using an automated platform (Ventana Discovery, Ventana Medical Systems; details of procedures can be obtained at request). The Fgf9 probe was a kind gift from D. Ornitz (Washington University), and the TrkC probe from L. Lei (UT Southwestern Medical Center, Texas, USA). All other probes were cloned from PCR fragments using the PCR-Script Amp cloning kit (Stratagene).

3.2.3 ChIP analysis

In silico analysis of the *Nt3* and *Fgf9* genes, including a 5,000-bp region upstream of the coding sequence, revealed several potential Zeb2-binding sites. Qualitative PCR primers were designed (Vector NTi) to cover 500-bp regions 4 kb upstream of the first exon. For *Fgf9*, we also included regions covering Zeb2 sites (fragments containing at least two CACCT(G) half-sites separated by less than 100 bp) in the introns. The performance of these primers was first validated on an AbiPrism qualitative PCR platform using a dilution series of genomic DNA. Later, these primers were used to assess the quality of the sheared or enzymatically digested chromatin (input DNA). To prepare the chromatin, we dissected mouse cortex and hippocampus tissue from E16.5–17.5 CD1 embryos in ice-cold Hank's solution, cut the tissue into smaller pieces, and dispersed by pipetting and counted the cells. The whole-cell suspension was then cross-linked in 1% formaldehyde (vol/vol) for 10 min at 24 °C. Cross-linking was stopped by addition of Glycine to a final concentration of 0.125 M. After several washes with PBS, the cell pellet was dissolved in lysis buffer (5 mM PIPES (pH8.0), 85 mM KCl and 0.5% NP40 (vol/vol) and kept on ice for 1-1.5 h to lyse the cells. The lysates were then passed 20 times through a 26G-needle to release the nuclei. Nuclei were lysed in nuclear lysis buffer (50 mM Tris-HCl (pH 8.1), 10 mM EDTA, 1% SDS (wt/vol) for 10–15 min before shearing. Chromatin was sonicated eight times for 30 s in a Branson Sonifier 450 (power setting 3) and with a constant duty cycle to obtain chromatin fragments of 500 bp (mean size). Samples were cooled for 1 min on ice between each pulse. An aliquot of the sheared chromatin was reverse-crosslinked and purified to check shearing efficiency and to measure DNA concentration. Samples were further processed according to the protocol provided with the ChIP-IT kit (Active Motif). For each immunoprecipitation reaction, 25 µg of chromatin was used. Immunoprecipitation was performed using 5 µg of crude anti-Zeb2 rabbit antibody or 5 µg of rabbit IgG as a negative control (Abcam). After recovery of the Zeb2-bound DNA fragments from the immunoprecipitation, fragments of *Nt3* and *Fgf9* promoters were detected by qualitative PCR on 20 ng of DNA (immunoprecipitate) per sample in triplicate, using the validated primers described above. For each region, at least three independent immunoprecipitation reactions were carried out. Next, the relative enrichment of the bound versus the unbound regions (after ChIP with antibody to Zeb2) was calculated relative to negative control rabbit IgG.

3.3 Results

In this project my aim was to (i) identify and validate direct target genes of Zeb2 during cortex development, which (ii) explain the underlying mechanism of the cortical phenotype in this series of conditional *Zeb2* KO mice. In addition, this was the first time in our laboratory we performed this on basis of global comparative gene expression profiling (at that time micro-array based, whereas it is now replaced by RNA-Seq).

We first carried out micro-array analysis of *Zeb2* KO versus (Zeb2) control (CTRL) forebrain tissue at E14.5 and E18.5. We selected the top-20 candidate genes (see Table. 1) that were upregulated in the *Zeb2* KO brains and therefore could be repressed by Zeb2.

List of potential target genes of Zeb2 in developing cortex				
	Gene	**Name**	**UP at E14.5**	**UP at E18.5**
1	Car7	carbonic anhydrase 7	yes	yes
2	BC049730	cDNA sequence BC049730	yes	no
3	Ap1m2	adaptor protein complex AP-1, mu 2 subunit	yes	yes
4	Nt3	neurotrophin 3	yes	yes
5	Trhr	thyrotropin releasing hormone receptor	yes	yes
6	Grm1	glutamate receptor, metabotropic 1	yes	yes
7	BC027061	hypothetical protein LOC232491	yes	no
8	Gpr88	G-protein coupled receptor 88	yes	yes
9	Gpr73l1	prokineticin receptor 2	yes	no
10	Dbpht2	DNA binding protein with his-thr domain	yes	no
11	Tesc	tescalin	yes	no
12	Rasgef1c	RasGEF domain family, member 1C	yes	no
13	Fgf9	fibroblast growth factor 9	yes	yes
14	Gm337	Mus musculus dual specificity phosphatase 5	yes	no
15	Ebf1	early B-cell factor 1	yes	no
16	Hrk	harakiri	yes	yes
17	AI427515	expressed sequence AI427515	yes	yes
18	Bmp2	bone morphogenetic protein 2	yes	yes
19	AK140273	hypothetical protein	yes	yes
20	Lad1	ladinin	yes	yes

Table 2. List of potential target genes of Zeb2 in developing cortex.

The list of transcripts downregulated upon *Zeb2* KO at E14.5 and E18.5 was significantly shorter and contained only four protein-encoding genes: Mc4R, Crk7, Efcbp1 and Ldb2. This indicated that Zeb2 indeed acts mostly as a transcriptional repressor.

My focus was on the validation of these potential target genes by RT-qPCR and *in situ* hybridization (ISH), to prioritize the list of candidates based on whether or not a candidate could confirm the obtained micro-array expression data. *Zeb2* is expressed in the upper layers of the cortex, but the phenotype (premature generation of the upper layers at the expense of the deep layers and premature

gliogenesis, see introduction to this Chapter) was pointing out towards defects in the progenitor zone where Zeb2 is absent. Because of this we focused on the secreted/diffusible factors whose expression could be (directly) regulated by Zeb2 in the upper layers and which could feed back to the progenitor zone to orchestrate timing of neurogenesis and numbers of neurons/glia generated.

Neurotrophin 3 (Nt3) and Fibroblast growth factor 9 (Fgf9) mRNA were strongly upregulated in the micro-array data set in the *Zeb2* KO versus Zeb2 CTRL samples. We subsequently found that Nt3 and Fgf9 transcripts were strongly over-expressed (RT-qPCR) in the mutant brains at E14.5 and 18.5, respectively (Fig.16 A,E). Nt3 and Fgf9 mRNA were detected in Zeb2-deficient cortical plate as early as E12.5. At E13.5, 14.5 and 15.5, Nt3 mRNA was markedly upregulated (even visible by ISH) when more post-mitotic cells began to populate the cortical plate (Fig.16 C, D and Fig. 17 B,D). Increased Fgf9 mRNA levels were detected prematurely at E16.5 in the mutant cortex (Fig. 16 G,H, Fig. 17 F).

To determine whether Fgf9 and Nt3 alone could mimic the *Zeb2* KO phenotype *in vitro*, we added these factors separately to the WT organotypic slice cultures of E16.5 brains. We could not reproduce the phenotype using Nt3. Notably, addition of Fgf9 (to the medium or as Fgf9-coated beads implanted in the E16.7 and E17.5 cortex slices) resulted in enhanced generation of Olig2+ glial precursors in the VZ and in the CP validating our Zeb2-dependent feedback theory (experiments done by A.N. and E.S., *data not shown here;* see Seuntjens *et al.,* 2009).

To verify if Fgf9 and Nt3 are direct targets of Zeb2, we carried out ChIP, using a polyclonal anti-Zeb2 antibody, on sheared chromatin isolated from cortical tissue at E16.5–17.5. Zeb2 was known to interact with regulatory elements containing a separated tandem of two sequences (CACCT(G) and/or, in fewer cases, CACANNT (Remacle *et al.*, 1999). We designed several primer pairs within 5 Kb upstream of the first exon of mouse *Nt3* and *Fgf9*, respectively, and analyzed the immunoprecipitate via RT-qPCR. We found a significantly higher ($P = 0.044$) level of interaction of Zeb2 with the fragment encompassing the −1,300 region of *Nt3*, but we could not detect any preferential binding of Zeb2 with the *Fgf9* upstream region. We also tested three intronic regions of *Fgf9* that contain putative Zeb2-binding sites, but none of these were preferentially enriched for Zeb2-binding.

Taken together, I showed that Zeb2 interacts directly with the promoter region of mouse *Nt3*, possibly downregulating its transcription (Fig.16 I). Although *Fgf9* was clearly dependent on intact Zeb2 levels, it remains unclear as to whether this is mediated via direct interaction of Zeb2 with a promoter or promoter-proximal regulatory element in *Fgf9* (Fig. 16 I). In order to respond to the growth factor cues a cell has to express the factor-specific receptors. So, we had also to investigate whether the expression domain and/or level of relevant receptor genes was altered. TrkC is the Nt3 receptor here, and FgfR2 is

the major candidate Fgf9 receptor here. We documented (by ISH) that both TrkC (E14.5) and FgfR2 (E16.5) were present in the proliferative zone of the cortex, but did not observe any differences, at least using ISH, between control and mutant brains (Fig. 16 j,m)

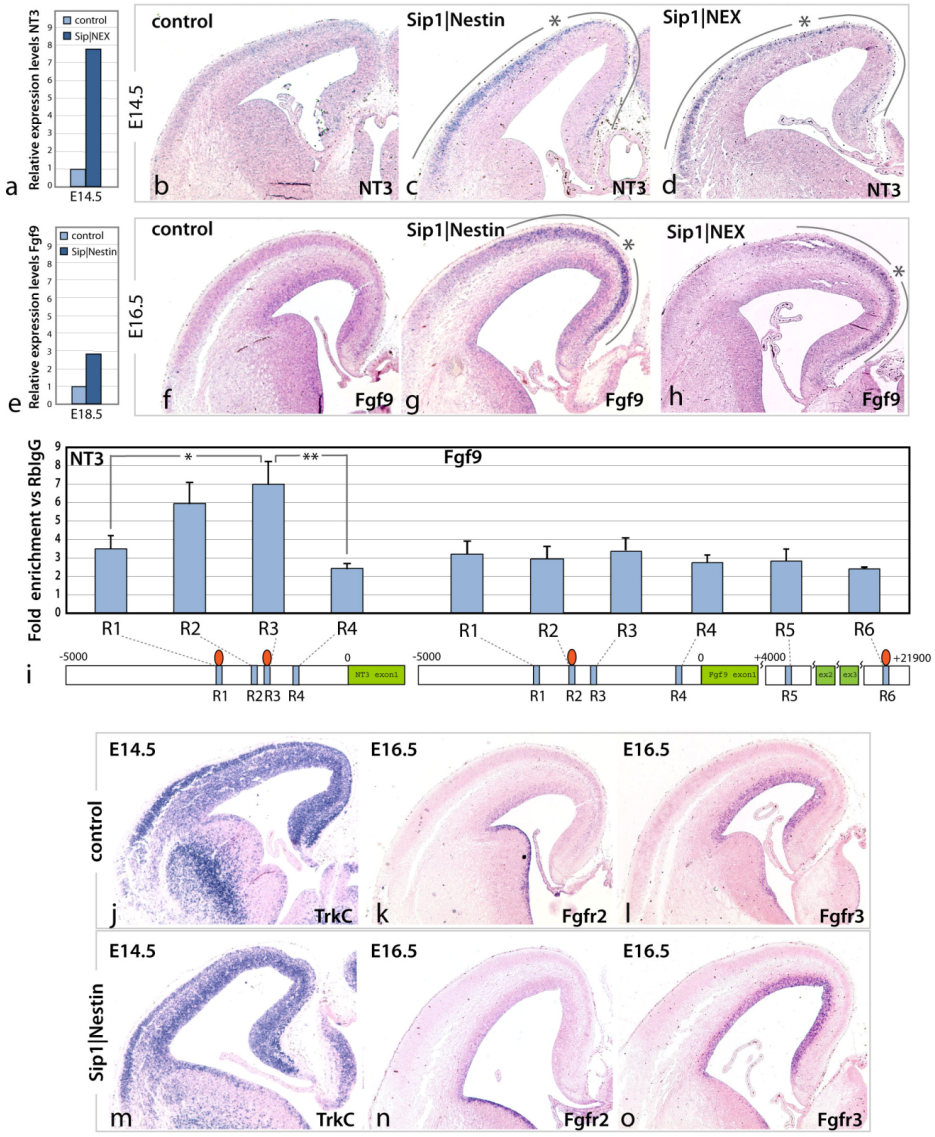

Figure 16. Expression *Nt3* and *Fgf9* and their receptors, verified by ISH, in developing cortex. ChIP analysis for Sip1 on Nt3 and Fgf9 promoters.

(a–f) ISH of Nt3 (a–c) and Fgf9 (d–f) mRNA at E14.5 and 16.5 respectively, showed upregulation of both in the Zeb2-deficient cortical plate. Fgf9 was found more in the cingulate cortex, whereas Nt3 was detected in the entire cortical plate (line with star). (g) Different regions (R) within 5-kb upstream of the first exon of the Nt3 and Fgf9 genes were tested for interaction with Sip1 by ChIP. The graph shows the relative enrichment of these regions for rabbit antibody to Zeb2–immunoprecipitated DNA versus rabbit IgG–immunoprecipitated DNA. In the sketch, gray rectangles depict the regions that were amplified by validated primer sets. Putative tandem Zeb2-binding sites are represented by red ellipses. Only R3 seems to be bound by Zeb2 (* P = 0.047 and** P = 0.044, error bars indicate s.e.m.). In the *Fgf9* regulatory region, we could not detect specific interaction with four regions in the upstream regulatory region and two regions in intronic sequences. (h,k) ISH for the Nt3 receptor TrkC at E14.5 showed high mRNA expression levels in the entire cortex. (i,l) ISH for the Fgf9 receptor Fgfr2 showed mRNA expression localized to the ventricular zone of the ganglionic eminences at E16.5. (j,m) *Fgfr3* at E16.5 was highly expressed in the cortical ventricular zone (Seuntjens *et al.*, 2009).

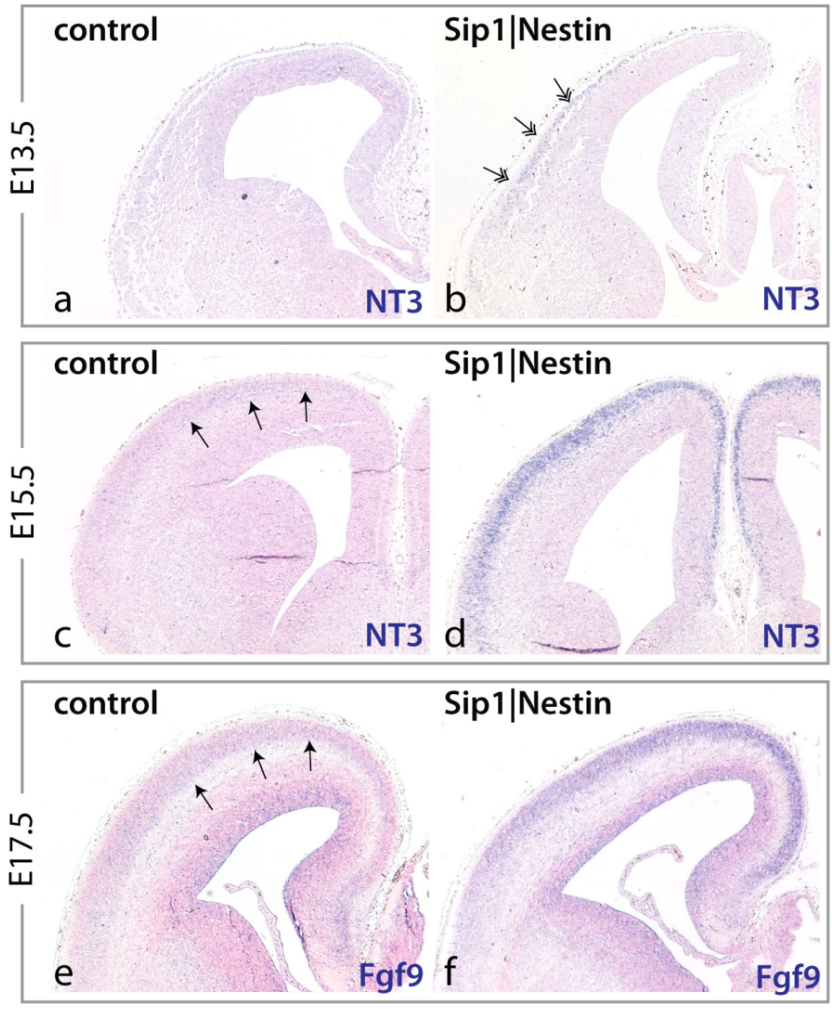

Figure 17. Nt3 and Fgf9 are expressed at low levels in developing cortex.

In situ hybridizations for NT3 at E13.5 (a-b), E15.5 (c-d) and for Fgf9 at E17.5 (e-f) show that their expression remains low in the control cortex (arrows). NT3 overexpression can already be detected at E13.5 (double arrows). (Seuntjens *et al.*, 2009)

3.4 Discussion

The fate of cortical progenitors that generate neurons and glial cells during development is determined by temporally and spatially regulated signaling mechanisms. We show that Sip1/Zeb2, which is present at high levels in post-mitotic neocortical neurons and cannot be detected in the precursor cells lining the embryonic (fore)brain ventricles, regulates progenitor fate. Furthermore, it does this in a non-autonomous fashion. Conditional genetic inactivation of *Zeb2* in induced premature production of upper-layer neurons at the expense of deep layers, precocious and increased generation of glial precursors,

and enhanced postnatal astrocytogenesis. Such premature upper-layer generation coincided with overexpression of *Nt3* and upregulation of *Fgf9* (and not of their respective cognate receptors) expression.

Nt3 was previously shown to play a role in neurogenesis. When cultured in presence of function-blocking Nt3 antibodies, cortical progenitors show decreased proliferation and survival what resulted in Mek/Erk-dependent inhibition of neurogenesis (Barnabé-Heider and Miller, 2003). Also, more recently, a similar study addressing the role of Nt3 in cortex development showed that injection of Nt3 along the VZ of E13.5 brains altered cell fate decisions and cell migration; neurons destined to become layer 4 switched now are part of upper layer production (layers 2/3). This phenotype was mimicked *in vitro* and could be rescued by inhibition of Mek/Erk signaling (Ohtsuka *et al.*, 2013). In a recent study Parthasarathy and co-workers (Parthasarathy *et al.*, 2014) showed that overproduction of Nt3 in the post-mitotic neurons at the onset of corticogenesis resulted in a switch from deep layers to upper layers cell generation. In contrast, loss of *Nt3* caused an increase in deep layer 6 neurons, but added Nt3 did not rescue the *Zeb2* KO phenotype suggesting other parallel mechanisms control the cell fate switches and the timing of neurogenesis. This data is in line with the phenotype observed in the *Zeb2* KO brains where overproduction of Nt3, seen at the mRNA level, leads to an enhanced generation of the upper layer neurons at the expense of the deep layers. Our data is further substantiated by precocious activation of Mek/Erk, downstream of neurotrophins (reviewed in Kaplan and Miller, 2000).

Fgf9 was previously implicated in the gliogenic switch in retinal pigment epithelium (Zhao *et al.*, 2001), and addition of Fgf9 to the Muller glial cell cultures (Muller glia are the main glial cell type of the retina) induced their proliferation (Cinaroglu *et al.*, 2005). Presently ongoing collaborative work between our lab and R. Ashery-Padan (Tel Aviv) on the role of Zeb2 in retinogenesis is addressing this as well. Recently, it was shown that Mek is a key regulator of gliogenesis in the developing brain. Mek1 and Mek2 are the core components of the Mek/Erk pathway here. It was shown that deletion of *Mek1/2* in the developing brain results in a reduction of glial progenitors whereas higher Mek1 levels promote precocious glial progenitor specification and termination of neurogenesis in a cell-autonomous way (Li *et al.*, 2012b).

This work shows that Zeb2 restrains the production of signaling factors in post-mitotic neurons that feed back to progenitors to regulate the timing of cell fate switch and the numbers of neurons and glial cells throughout corticogenesis. It could be that in absence of Zeb2, *Nt3* is derepressed in the upper layers of the cortex, which causes precocious activation of the Mek/Erk pathway. Precociously activated Mek/Erk signaling subsequently induces premature glial progenitor specification program enhanced by Fgf9 overproduction. Thus, *Fgf9* upregulation and induced gliogenesis could be secondary to the *Nt3* induction and could be triggered by Mek/Erk pathway activation.

It has previously been shown that extrinsic factors from CSF, blood vessels, meninges or produced locally within the VZ of the cortex influence cell fate decisions during development (see Introduction section to this Chapter). We have shown, for the first time in the field, that a transcription factor (Zeb2) acts non-autonomously in the post-mitotic cells of the brain cortex and, doing so, regulates the timing of neurogenesis and gliogenesis in the remote progenitor zone of the cortex.

4 Chapter 4: Role(s) of Zeb2 in pluripotency and differentiation of embryonic stem cells

Previously published studies on the role of Zeb2 in early cell fate decisions in human ESCs and mouse EpiSCs using shRNA-mediated knockdown of Zeb2 and in parallel overproduction of Zeb2 showed that it is an important regulator or ectodermal vs. mesendodermal fate (Chng *et al.*, 2010).

Our laboratory was the first to derive and study *Zeb2* knockout ESCs. We hypothesized that permanent genetic inactivation of *Zeb2* may result in a different, possibly more severe phenotype than knockdown, allowing us to better understand Zeb2's mechanisms of action that otherwise were masked by incomplete silencing of Zeb2. The first goal of this part of the project was to develop a convenient and simple, robust differentiation system(s) for mouse ESCs that would allow us to get insight into how Zeb2 regulates cell fate decisions starting from the ground pluripotent state. Because of the well-documented link between *ZEB2* gene mutations and severe and milder forms of Mowat-Wilson syndrome, the emphasis was put on developing an efficient neural differentiation protocol from pluripotent cells. The second goal was to document - at mRNA and DNA-methylation level - and to further study the consequences of *Zeb*2 genetic inactivation in ESCs both for their pluripotent state as well as for general and directed neural differentiation, respectively (for protocols see Methodology and Materials section below). Because of the observed *Zeb2* KO phenotype, my focus was on the relation between Zeb2, pluripotency network and enzymes regulating DNA-methylation status in ESC pluripotency and during differentiation. The aforementioned part of my work (see also Chapters 4.1, 4.3.1- 4.3.7, 4.4.1) is currently under revision. We also explored additional mechanisms by which Zeb2 could regulate ESC differentiation and doing so we identified the neuronal-inhibitory factor REST) which forms multiple complexes) as downstream target of Zeb2 (Chapter 4.3.8).

In addition we performed teratoma formation assays and we discovered that *Zeb2* KO ESCs injected into immunodeficient mice could generate teratomas comprised of cells of either of the three germ layers (Chapter 4.3.9). The results from the teratoma experiments prompted us to study cell-autonomous vs. cell non-autonomous action of Zeb2 *in vitro*. For this we performed co-culture experiments where we mixed labeled Ctrl and *Zeb2* KO ESCs in different ratios, subjected them to neural and general differentiation, respectively, and assessed whether Ctrl cells were able to direct in this set-up the differentiation of *Zeb2* KO ESCs (Chapters 4.3.10 and 4.4.3). To test the *in vivo* developmental potential of Zeb2 KO ESCs, we performed chimera experiments (Chapter 4.3.11).

We also asked whether interaction between Zeb2 and its Smad partners would be beneficial or inhibitory for neural differentiation of ESCs. For this we generated a series of ESC lines where we inserted wild-type and (the appropriate) mutant Zeb2 cDNAs in the R26 locus of the Zeb2 KO ESCs (Chapters 4.3.12 and 4.4.4).

4.1 Introduction

This introduction specifically addresses the main part of my ESC-related project (Chapters 4.3.1- 4.3.7, 4.4.1) and is currently under revision (Stryjewska, Dries *et al.*).

Naïve mouse embryonic stem cells (mESCs) (Evans and Kaufman, 1981; Martin, 1981), primed epiblast stem cells (EpiSCs) (Brons *et al.*, 2007; Tesar *et al.*, 2007) and embryonic germ cells (Matsui *et al.*, 1992; Leitch *et al.*, 2010) are pluripotent cells that can be propagated *in vitro* and used as cell models to study pluripotent cell states and subsequent fate decisions that occur at different stages of embryonic development, transitions that require changes of the transcriptome and methylome. For cultured populations of ESCs, both adherent cells and embryoid bodies (EBs) are used. The ground-state (more stable than the metastable, naïve state) of self-renewing mESCs can be achieved by simultaneous addition of LIF and chemical inhibitors (referred to as 2i) of MAPK and GSK3 signaling (Ying *et al.*, 2008). When compared to a population of naïve ESCs, ground-state ESCs are characterized by higher and more homogeneous expression of key pluripotency genes, lower levels of differentiation markers and reduced DNA-methylation (Marks *et al.*, 2012; Leitch *et al.*, 2013).

DNA-methylation status has profound effects on embryonic gene expression. It is controlled by DNA (cytosine-5)-methyltransferases (Dnmts; Dnmt3a/3b/3l) that are highly active in ESCs and early embryos and establish new methylation patterns, and by Dnmt1 (the maintenance methyltransferase) that copies the patterns onto daughter cells (Okano *et al.*, 1999; Goll and Bestor, 2005). Active demethylation is orchestrated by Ten-eleven translocation (Tet) methylcytosine dioxygenases (Tahiliani *et al.*, 2009; Kohli and Zhang, 2013). Tet1 levels are high in ESCs and decrease upon differentiation, correlating with exit from pluripotency, and Tet1 steers mesendoderm versus trophectoderm decisions in the pre-implantation embryo (Ito *et al.*, 2010; Koh *et al.*, 2011). Tet1 is also important during somatic reprogramming for genome demethylation as well as activation and maintenance of Oct4/Nanog functions (Takahashi and Yamanaka, 2006; Costa *et al.*, 2013; Gao *et al.*, 2014).

The transcription factor Zeb2 (Sip1, Zfhx1b) downregulates *E-cadherin* (*Cdh1*) and thereby contributes to EMT (Comijn *et al.*, 2001), which is relevant to stem cell fate but also to tumorigenesis (Kim *et al.*, 2014; Pieters and van Roy, 2014). Mutations in *ZEB2* cause Mowat-Wilson syndrome (OMIM #235730), including defects in the central and peripheral nervous system (CNS, PNS) (Mowat *et*

al., 1998; Cacheux *et al.*, 2001; Wakamatsu *et al.*, 2001). Many *in vivo* studies confirm the critical roles of Zeb2 in embryogenesis and neurodevelopment in particular. *Zeb2* KO mouse embryos die shortly after E8.5 and have multiple defects including in somitogenesis (Maruhashi *et al.*, 2005), the neural plate and its flanking, future neural crest cells (Van de Putte *et al.*, 2003). Cell-type specific *Zeb2* KO mice develop defects in the CNS (Seuntjens *et al.*, 2009; McKinsey *et al.*, 2013; van den Berghe *et al.*, 2013) and PNS (Van de Putte *et al.*, 2007; Jeub *et al.*, 2011; Weng *et al.*, 2012). Such studies in the embryonic brain cortex revealed cell-autonomous, but also non-autonomous Zeb2 actions. In human (h) ESCs, Zeb2 regulates cell fate: upon Zeb2 KD they commit towards mesendoderm, while Zeb2 overproduction leads to enhanced neurogenesis (Chng *et al.*, 2010). *ZEB2* is controlled by Nanog, Oct4 and Sox2 in hESCs, but key genes downstream of Zeb2 in m/hESCs, and during early neural development, remain to be determined, and *Zeb2* KO hESCs have not been reported yet.

In order to enter lineage commitment, the pluripotency network in ESCs and EpiSCs needs to be distinguished (Festuccia *et al.*, 2013; Trott and Martinez Arias, 2013). The list of factors promoting exit from naïve or ground state is growing, yet more key players remain to be identified (Kaji *et al.*, 2006; Betschinger *et al.*, 2013; Leeb *et al.*, 2014). Exit from pluripotency beyond the primed epiblast state requires efficient, irreversible silencing of the transcriptional pluripotency network (including *Oct4* and *Nanog* silencing, which persist in EpiSCs), acquisition and maintenance of DNA-methyl marks, and initiation of differentiation. Using *Zeb2* KO cells we identified Zeb2 as a new, critical player in ESCs to initiate and execute their three-lineage differentiation program. Upon withdrawal of LIF+2i, *Zeb2* KO ESCs only sometimes commit to differentiation, instead they usually stall as a pluripotent, epiblast-like cell population that maintains the ability to re-adapt to 2i even after prolonged exposure to differentiation protocols. The defective silencing of the pluripotency program prevents these *Zeb2* KO cells from undergoing neural and general (including mesendodermal) differentiation. RNA-seq revealed that Dnmt and Tet family mRNA levels are deregulated in *Zeb2* KO cells. *Zeb2* KO cells correctly acquire methyl marks early during neural differentiation, but do not maintain these and revert to a more naïve methylome state. Tet1 levels depend on the presence of Zeb2 and in *Zeb2* KO cells (displaying elevated Tet1) Tet1 KD rescues their ability to exit from their pluripotent state and re-enter lineage commitment.

4.2 Methodology and Materials

4.2.1 ESC lines

All experiments on live mice used for deriving the embryos for establishing the ESC lines were performed in accordance with relevant institutional (KU Leuven approved project P153/2012), national (lab license LA1210584 issued by the Belgian federal government) an international (directive 2010/63/EU) guidelines and regulations. The KU Leuven Ethical Committee approved the experiments and confirmed that all experiments were done conform to the relevant regulatory standards.

Two independent ESC derivations were performed. First, control lines were derived by interbreeding *Zeb2*$^{flox/flox}$ mice (CD1 background: Higashi *et al.*, 2002). From these ESC lines, *Zeb2* KO lines were derived by transient transfection of vector-encoded Cre. Five control ESC lines and two KO lines were established. Secondly, *Zeb2*$^{+/-}$ mice were crossed with R26-iPSC mice that contain an RMCE cassette in the ROSA26 locus (Haenebalcke *et al.*, 2013a). The second ROSA26 allele contained the LacZ reporter (Soriano, 1999). New control and RMCE-compatible *Zeb2* KO ESC lines (3 clones; mixed 129/Bl6 background) were derived using a protocol (Pieters *et al.*, 2012) in which pluripotin was replaced with 1 µM PD0325901 (Axon, 1408) and 3 µM CHIR99021 (Axon, 1386). To obtain R26_Zeb2 lines, RMCE technology (Haenebalcke *et al.*, 2013b) was used to insert N-terminally flag-tagged, wild-type Zeb2 cDNA into the Rosa26 locus of the *Zeb2* KO ESCs.

4.2.2 ESC maintenance

ESCs were maintained feeder-cell free in 2i+LIF. N2B27 was prepared as described (Gaspard *et al.*, 2009). For 2i+LIF medium, 0.4 µM PD0325901 (Axon, 1408) was added together with 1 µM CHIR99021 (Axon, 1386; 1 µM) (Ying *et al.*, 2008). LIF (Millipore, ESG1107) was added to a final concentration of 1,000 U/ml.

4.2.3 Neural differentiation

On d0, 3x10^6 ESCs were plated in a 10-cm bacterial petri dish in EB medium (KO DMEM (Invitrogen, 10829018), 15% FBS (Life Technologies, 10270106), 0.1 mM non-essential amino acids, 1 mM sodium pyruvate, 0.1mM 2-mercaptoethanol, 50 U/ml of penicilline/streptomycine). On d2 the EB medium was refreshed. On d4 the medium was changed to N2B27 medium and retinoic acid (Sigma, R2625) was added to a final concentration of 500 nM. On d6 the medium was refreshed. Between d8 and d15 the EBs were cultured in N2B27, and this medium was refreshed every other day

4.2.4 General differentiation

On d0, $3x10^6$ ESCs were plated in a 10-cm bacterial petri dish in EB medium (KO DMEM (Invitrogen, 10829018), 10% FBS (Life Technologies, 10270106), 0.1 mM non-essential amino acids, 1 mM sodium pyruvate, 0.1mM 2-mercaptoethanol, 50 U/ml of penicilline/streptomycine). EB medium was changed every other day till d15.

4.2.5 EB dissociation and sorting of living cells

EBs on d15 were dissociated using Liberase (Roche, 05401020001). Living cells were stained with propidium iodide (Sigma-Aldrich, P4864) shortly before sorting.

4.2.6 Teratoma formation assay

For teratoma formation assay Rag2c -/-;Gamma(c) -/- mice were injected $1x10^6$ ESCs in 50µl PBS. Four mice were injected with Zeb2 KO ESCs (clone 13 and clone 18) and 4 mice were injected with Zeb2 Ctrl ESCs (clone 2 and clone1). All injected ESC clones gave rise to teratomas. The animals were sacrificed 5-6 weeks after injection.

4.2.7 Morula aggregations

Morula aggregations were performed by Zhiyong Zhang from InfraMouse (KU Leuven).

Briefly, E2.5 morulae were isolated from pregnant superovulated CD1 females and the zona pellucida was removed. ESCs colonies were dissociated into small, 5-15-cell clumps using Accutase and resuspended in 2i medium prior to aggregation. Next, morulae and ESCs were transferred into a single drop of KSOM medium and aggregated overnight in a cell culture incubator (37°C, 5% CO_2). The following day blastocysts and compacted morulae were transferred into pseudopregnant female mice (10-20 blastocysts/mouse).The chimeric embryos were analyzed on E9.5.

4.2.8 RNA-sequencing and data analysis

Total RNA isolation was performed using a Qiagen RNeasy mini kit (Qiagen, 74104) according to the manufacturer's instructions. cDNA libraries were generated with Truseq RNA kit and sequenced according to the Illumina TruSeq v3 protocol on the HiSeq2000 with a single read 36 bp and 7bp index. Sequenced read fragments were mapped to the mouse genome assembly GRCm38 (Ensembl) using Tophat2 (v2.0.13). A count table for annotated genes was generated using featureCounts (v1.4.6). Genes were further classified in different biotypes based on Vega gene and transcript annotation (http://vega.sanger.ac.uk/info/about/gene_and_transcript_types.html).

RNAseq expression data: To compare counts between samples we converted them to Transcript Per Million (TPM) values. To retain only informative genes we filtered based on biotype, expression and standard variability using aforementioned TPM values. First we removed all genes that belong to short non-coding categories, in the next step we selected only these genes that have at least 5 transcripts/million in minimum 3 samples and finally we removed the 20% lowest variable genes. The raw counts were imported in the R Package DESeq2 (Love *et al.*, 2014) to test for differential expression between pairwise time-points of KO and Ctrl samples.

RNAseq clustering: All samples were clustered using Principal Component Analysis (PCA) or unsupervised hierarchical clustering based on 1 – Spearman correlation distance scores with average linkage.

RNAseq gene ontology: To identify biological processes that are negatively enriched in Zeb2 KO, we sorted genes according to their pi-value (-log10(q-value) * logFC) based on DESeq2 timeseries analysis. The obtained ranked list was used as input for the GseaPreranked tool with only parameters *–nperm 3000, -set_max 500 -set_min 10* deviating from the default parameters.

RNA-seq motif sequence analysis: To perform simple motif sequence analysis between KO and Ctrl at d6 we defined promoter regions as ± 2Kb from the transcription start site (TSS) and counted the occurrences for the putative binding site of Zeb2 (double YACCTG sequences with maximum gap of 40 bases) for all (up and down) DEGs ($p<0.01$ and absolute log2 FC > 1); and, as background, the promoter regions of all genes. One-sided Fisher's exact test was used to determine significant over or underrepresentation of this motif in promoter regions of DEGs relative to the genome wide promoter regions.

4.2.9 RRBS analysis

Total genomic DNA was isolated by proteinase K digestion and precipitation with isopropanol. RRBS was performed by NXT-Dx (http://www.nxt-dx.com/) using the premium RRBS kit (Diagenode) according to the manufacturer's instructions.

RRBS processing: Quality of sequencing reads was first assessed by FastQC (v0.11.3_devel) and Trim Galore (v0.3.7) in *–rrbs* mode. These reads were then mapped to mouse genome GRCm38 (Ensembl) using bismark (v0.14.1) with parameters *–bowtie2 –maxins* 1000, allowing a maximum insert size of 1000 bp for paired-end sequences. To extract methylation information in a CpG context from both strands we used bismark_methylation_extractor with parameters *–paired-end –no_overlap –comprehensive*. We used the R package methylKit (Akalin *et al.*, 2012) and custom R scripts to further analyze the data. In brief, we considered only CpGs with a minimum sequencing depth of 5x and removed the top 0.1% with highest coverage. To visualize global percentage methylation histograms were created with 5%-methylation bins. For all further analyses we only retained CpGs that were present in all samples.

RRBS genomic regions: Genomic coordinates for genes were retrieved from the GRCm38 (Ensembl) annotation and only coordinates for protein-coding genes were used. We downloaded mm9 enhancer coordinates provided at http://chromosome.sdsc.edu/mouse/download.html ,converted them to mm10 coordinates using CrossMap (v0.1.8) and extended them in both directions with 1kb. CpG islands (CGI) and transposable elements (TE) were retrieved via the UCSC table browser for the GRCm38/mm10 genome, with respectively the CpG Islands and RepeatMasker track. The genomic coordinates for Canyons were retrieved from a published dataset (Jeong *et al.*, 2014a). We used a CpG observed/expected ratio of 0.29 to distinguish low- and high-CpG density promoters as described (Etchegaray *et al.*, 2015). Regions that do not belong to any of the aforementioned regions (*e.g.*. intergenic space) are described as "other" for simplification. To annotate demethylated enhancers to topological associated domains (TADs) we downloaded mouse ESC specific topological domains provided at http://chromosome.sdsc.edu/mouse/hi-c/download.html. We selectively used the combined *Hin*dIII dataset after converting to mm10 coordinates as described before.

RRBseq data analysis: To identify differentially methylated regions (DMRs) and analyze global methylation dynamics and differences we averaged methylation in 400bp tiles containing at least three CpGs. Tiles with more than 20% difference in methylation and a q-value < 0.05 were assigned as significant DMRs, or simply DMRs.

4.2.10 Analysis of published Tet1 binding peaks in mESCs

Data for Tet1 ChIP-seq for mouse ESCs was downloaded from GEO (GSM659803 and GSM659799). Reads were aligned to GRCm38 using bowtie with parameters *–e 70 –k 1 –m 1 –n 2 –concise*. Peaks were called with MACS software using default parameters. To study enrichment of Tet1 at demethylated regions, peaks were assigned to the closest demethylated region.

4.2.11 shRNA-mediated knockdown

Control and Tet1 lentiviruses were produced by standard methods:
(https://www.addgene.org/tools/protocols/pLKO/).
Control shRNA was used by combining MISSION Target shRNA in control vector SHC002 (Sigma).
The Tet1 shRNA (shTet: 5'-tcatctacttctcacctagtg-3') was cloned into pLKO1 vector.
REST knockdown line was generated by simultaneous infection with two RESTshRNA-encoding lentiviruses (TRCN0000321488 and TRCN0000071345) from the Sigma Mission Library.

4.2.12 ChIP analysis

$10x10^6$ ESCs (R26_Zeb2) were used per experiment. Cells were cross-linked for 10 min with ice-cold 1% formaldehyde. Cross-linked ESCs were sonicated using Branson Digital Sonifier (10 pulses, 30 sec

ON;60 sec OFF, amplitude 10). 10µg of Flag (Sigma, F3165) and 10µg of control mouse IgG (Santa Cruz, sc-2025) were used. Chromatin isolation and ChIP were done as described (Lee *et al.*, 2006). Phenol-chloroform purified DNA was used as template for qPCR to amplify proximal promoters of Nanog and E-cadherin. Primer sequences are listed in Table 1, section 4.2.11.

4.2.13 Immunohistochemistry and immunofluorescence

EBs were fixed overnight with 4% paraformaldehyde in PBS followed by progressive alcohol-assisted dehydration and paraffin embedding. 6-µm thick sections were used for IHC and IF. IHC and IF analyses were carried out on the Ventana Ultra Discovery platform (Roche). The following antibodies against the indicated proteins were used: Zeb2 (custom antibody; Seuntjens *et al.*, 2009), BIII-Tubulin (Abcam, ab78078), Oct4 (Abcam, ab19857), Nanog (Abcam, 80892), E-cadherin (BD Transduction Laboratories, 610182), Tet1 (Millipore, 09-872) and Alexa Fluor–tagged secondary antibodies (Jackson Immunoresearch) .

ESCs were fixed for 10 min with ice-cold 4% paraformaldehyde in PBS. Cells were blocked for 30 min at 24°C with 0.1% Triton X100-1% BSA in PBS. Anti-Oct4 (Abcam, ab19857;1/1000) and anti-Nanog (Abcam, 80892; 1:1000) were used as primary antibodies and DAPI was used as a nuclear counterstain (Life Technologies, D1306). Alexa Fluor–tagged secondary antibodies (Jackson Immunoresearch) were used at a dilution of 1:1000.

4.2.14 RT-qPCR primers

RT-qPCR primers:		
Gene	**Orientation**	**Sequence**
Afp	Fwd	CTTCCCTCATCCTCCTGCTAC
Afp	Rev	ACAAACTGGGTAAAGGTGATGG
Cdx2	Fwd	CAAGGACGTGAGCATGTATCC
Cdx2	Rev	GTAACCACCGTAGTCCGGGTA
Cer1	Fwd	CAACCACGAGGAGGCAGAAG
Cer1	Rev	GATCGCTTTCCACATCCCTT
Chrd	Fwd	CTGCGCTCAAGTTTACGCTTC
Chrd	Rev	AGGGTGTTCAAACAGGATGTTG
Cxcr4	Fwd	GACTGGCATAGTCGGCAATG
Cxcr4	Rev	AGAAGGGGAGTGTGATGACAAA
Dnmt3a	Fwd	GAGGGAACTGAGACCCCAC
Dnmt3a	Rev	CTGGAAGGTGAGTCTTGGCA
Dnmt3b	Fwd	AGCGGGTATGAGGAGTGCAT
Dnmt3b	Rev	GGGAGCATCCTTCGTGTCTG

Eomes	Fwd	CCTGGTGGTGTTTTGTTGTG
Eomes	Rev	TTTAATAGCACCGGGCACTC
Esrrb	Fwd	TTAACGCCATCCCCAAGCGCC
Esrrb	Rev	CAAGGCGCACACCTTCCTTCAGC
Fillagrin	Fwd	ATGTCCGCTCTCCTGGAAAG
Fillagrin	Rev	TGGATTCTTCAAGACTGCCTGTA
Foxa2	Fwd	TCCGACTGGAGCAGCTACTAC
Foxa2	Rev	GCGCCCACATAGGATGACA
Gata4	Fwd	CACCCCAATCTCGATATGTTTGA
Gata4	Rev	GCACAGGTAGTGTCCCGTC
Gata6	Fwd	TTGCTCCGGTAACAGCAGTG
Gata6	Rev	GTGGTCGCTTGTGTAGAAGGA
Gfap	Fwd	CCAGATCCGAGAAACCAGCCTGGA
Gfap	Rev	TGAGGTGGCCTTCTGACACGGA
Gsc	Fwd	CCCCGGTTCTGTACTGGTG
Gsc	Rev	TCTGGGTACTTCGTCTCCTGG
Hnf4	Fwd	CACGCGGAGGTCAAGCTAC
Hnf4	Rev	CCCAGAGATGGGAGAGGTGAT
Krt1	Fwd	TGGGAGATTTTCAGGAGGAGG
Krt1	Rev	GCCACACTCTTGGAGATGCTC
Krt10	Fwd	CGAAGAGCTGGCCTACCTAAA
Krt10	Rev	GGGCAGCGTTCATTTCCAC
Krt14	Fwd	GAGGAGACCAAAGGCCGTTAC
Krt14	Rev	GAGGAGAATTGAGAGGATGAGGA
Krt18	Fwd	GTCAGAGACTGGGGCCACTA
Krt18	Rev	CTCTAAAGTCATCGGCGGCAA
Lefty1	Fwd	CCAACCGCACTGCCCTTAT
Lefty1	Rev	CGCGAAACGAACCAACTTGT
Lefty2	Fwd	CAGCCAGAATTTTCGAGAGGT
Lefty2	Rev	CAGTGCGATTGGAGCCATC
Map2	Fwd	GCCAGCCTCAGAACAAACAG
Map2	Rev	AAGGTCTTGGGAGGGAAGAAC
Mixl1	Fwd	ACGCAGTGCTTTCCAAACC
Mixl1	Rev	CCCGCAAGTGGATGTCTGG
Nanog	Fwd	TCTTCCTGGTCCCCACAGTTT
Nanog	Rev	GCAAGAATAGTTCTCGGGATGAA
Nodal	Fwd	TTCAAGCCTGTTGGGCTCTAC
Nodal	Rev	TCCGGTCACGTCCACATCTT
Oct-04	Fwd	AGAGGATCACCTTGGGGTACA
Oct-04	Rev	CGAAGCGACAGATGGTGGTC
Pax7	Fwd	TCTCCAAGATTCTGTGCCGAT

Pax7	Rev	CGGGGTTCTCTCTCTTATACTCC
Pdgfra	Fwd	TCCATGCTAGACTCAGAAGTCA
Pdgfra	Rev	TCCCGGTGGACACAATTTTTC
Pdgfrb	Fwd	TTCCAGGAGTGATACCAGCTT
Pdgfrb	Rev	AGGGGGCGTGATGACTAGG
Prdm14	Fwd	CTCTTGATGCTTTTCGGATGACT
Prdm14	Rev	GTGACAATTTGTACCAGGGCA
Sox17	Fwd	GATGCGGGATACGCCAGTG
Sox17	Rev	CCACCACCTCGCCTTTCAC
T	Fwd	CTCGGATTCACATCGTGAGAG
T	Rev	AAGGCTTTAGCAAATGGGTTGTA
Tet1	Fwd	ACACAGTGGTGCTAATGCAG
Tet1	Rev	AGCATGAACGGGAGAATCGG
Tet2	Fwd	AGAGAAGACAATCGAGAAGTCGG
Tet2	Rev	CCTTCCGTACTCCCAAACTCAT
Tet3	Fwd	TCTCTGAAGGGTGGATTGTCC
Tet3	Rev	CCCAGCACCGAGTAGCTTTC
VegfR2	Fwd	TTTGGCAAATACAACCCTTCAGA
VegfR2	Rev	GCAGAAGATACTGTCACCACC
Zeb1	Fwd	ACCGCCGTCATTTATCCTGAG
Zeb1	Rev	CATCTGGTGTTCCGTTTTCATCA
Zeb2	Fwd	ACCTTACGAATGCCCAAACTGCA
Zeb2	Rev	ACAGAATTAGGGGAAGAACCCGTCT
Zfp42	Fwd	CCCTCGACAGACTGACCCTAA
Zfp42	Rev	TCGGGGCTAATCTCACTTTCAT
Zfp521	Fwd	GAGCGAAGAGGAGTTTTTGG
Zfp521	Rev	AGTTCCAAGGTGGAGGTCAC
RT-qPCR ChIP primers:		
Gene	**Orientation**	**Sequence**
Cdh1 Ctrl region	Fwd	TGAGGTCCTAGGTTCCATCTC
Cdh1 Ctrl region	Rev	GAAGGCAGGAACTGAACACA
Cdh R1	Fwd	GCTAGGCTAGGATTCGAACGAC
Cdh R1	Rev	TGCAGGGCCCTCAACTT
Nanog R1	Fwd	CAGCCGTGGTTAAAAGATGAATAAAGTG
Nanog R1	Rev	CAGCCGTGGTTAAAAGATGAATAAAGTG
Nanog Ctrl region	Fwd	GGTGATACGTTGGCCTTCTAGTCTGAA
Nanog Ctrl region	Rev	GGGCAAATTGCAAACTAACTGTATAACCTC

Table 3. RT-qPCR primers.

4.3 Results

4.3.1 Knockout of *Zeb2* impairs ESC differentiation

We generated *Zeb2* KO mESCs (deleting critical exon7, Higashi *et al.*, 2002; Van de Putte *et al.*, 2003) along with control ESCs (Ctrl; with exon7 floxed in both Zeb2 alleles). In 2i medium, Ctrl and *Zeb2* KO ESC populations maintain high Nanog and Oct4 (Fig. 19A), proliferate at comparable rates (Fig. 19B) and have a high and similar clonogenic capacity (±70%, *data not shown*), showing that Zeb2 is dispensable for pluripotency and self-renewal in ground-state conditions. Because of its known role in neural development (Mowat *et al.*, 1998; Miquelajauregui *et al.*, 2007; Seuntjens *et al.*, 2009; Chng *et al.*, 2010; van den Berghe *et al.*, 2013) we investigated the neural differentiation (ND) potential of *Zeb2* KO ESCs, subjecting the ESCs as embryoid bodies (EBs) to ND using retinoic acid (modified from Bain *et al.*, 1996) (Fig. 18A). In Ctrl EBs the very low Zeb2 mRNA level increased between day (d) 0 and d4 after 2i withdrawal as well as during the acquisition of neural fate (between d4 and d6) and remained high till the end of our 15-day long ND protocol (Fig. 18B). The first Zeb2-positive (Zeb2+) cells are detected by immunohistochemistry (IHC) on d6, being intense from d8 (*not shown*) till the end of the experiment (Fig. 18C). Absence of neural progenitor (Nestin), neuronal (BIII-Tubulin, Map2) and astroglial (GFAP) markers by immunofluorescence (IF) (Fig. 18D-F) showed that ND was abolished in *Zeb2* KO EBs. Thus Zeb2 is crucial for mESCs to acquire neural fate, in line with previous observations that Zeb2 KD in hESCs makes these cells favor mesendodermal over neuroectodermal fate (Chng *et al.*, 2010). To validate whether genetic inactivation of *Zeb2* in mESCs would also result in increased mesendoderm formation, we subjected the *Zeb2* KO ESC to a general differentiation (GD) protocol allowing commitment to all cell fates for 15 days, and monitored Zeb2 mRNA/protein in the Ctrl cells (Fig. 19C-E), and stained for mesodermal, endodermal and neural markers, respectively (Fig. 19F-H). This showed that *Zeb2* KO mESC have an impaired early differentiation that is not restricted to ND, but affects all three germ layers.

Gene expression changes in *Zeb2* KO mESCs after exposure of the cells to differentiating cues were also analyzed documenting 40 marker mRNAs for neuroectoderm, mesoderm, endoderm and pluripotency, respectively, by RT-qPCR on d0, 4, 6 and 15 in Ctrl and *Zeb2* KO cells, in ND and GD. Zeb2 "rescue" ESC lines were included in this analysis (d0 and d15; Fig. 20F,G) by introducing Zeb2 (N-terminally tagged with $Flag_3$/Strep-tag) in the *R26* locus (see Supplemental Experimental Procedures) of *Zeb2* KO cells (hereafter named R26_Zeb2), which restored differentiation (for IHC/IF and RT-qPCR, see Fig. 20 A-E). The expression heatmap (Fig. 18G) with samples clustered based on quantile-normalized expression values showed clear separations between d15 Ctrl and R26_Zeb2 cells

both in GD and ND, the d6 Ctrl in ND, and the rest of the samples including d15 *Zeb2* KO cells, further supporting our initial observation that *Zeb2* KO ESCs stay largely uncommitted and display overall reduced differentiation capacity.

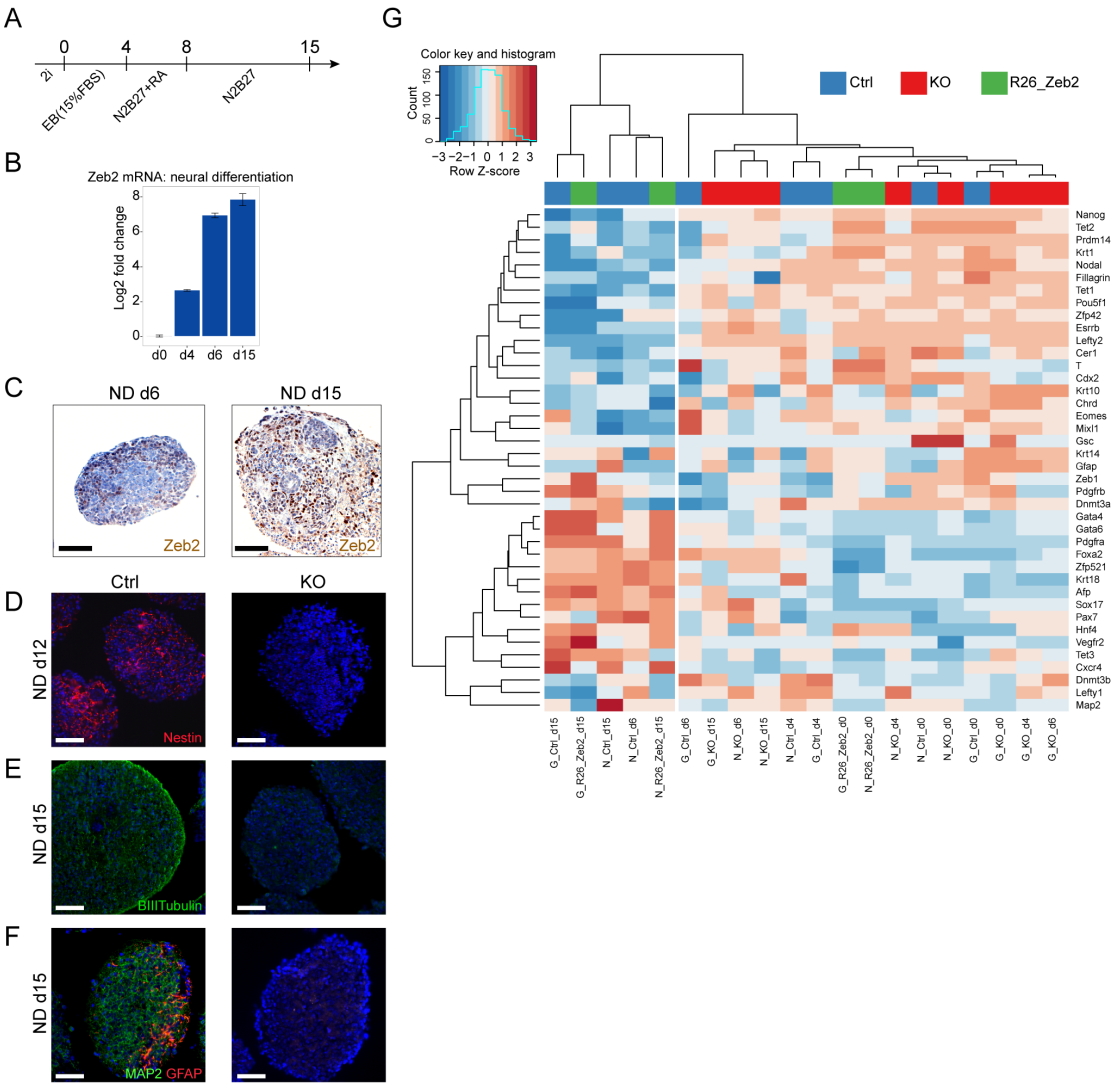

Figure 18. Knockout of Zeb2 impairs ESC neural differenitation.

(for general differentiation, see Fig. 19). A. Scheme of the 15-day ND protocol. B. RT-qPCR of Zeb2 in Ctrl ESCs during ND. SD of 2 technical replicates is shown. C. IHC for Zeb2 (brown) in Ctrl embryoid bodies (EBs) (Ctrl) on d6 and d15 of ND. D-F. Ctrl and *Zeb2* KO (KO) ND-EBs stained for Nestin (red, panel D) on d12, BIII-Tubulin (green, panel E) on d15 and co-stained for MAP2 (green) and GFAP (red) on d15 (panel F). Scale bars: 50µm. Results shown are from one experiment and are representative for 3 experiments. G. Heatmap for samples collected in pluripotency and during ND and GD with clustering based on Spearman correlation distances of quantile-normalized RT-qPCR values.

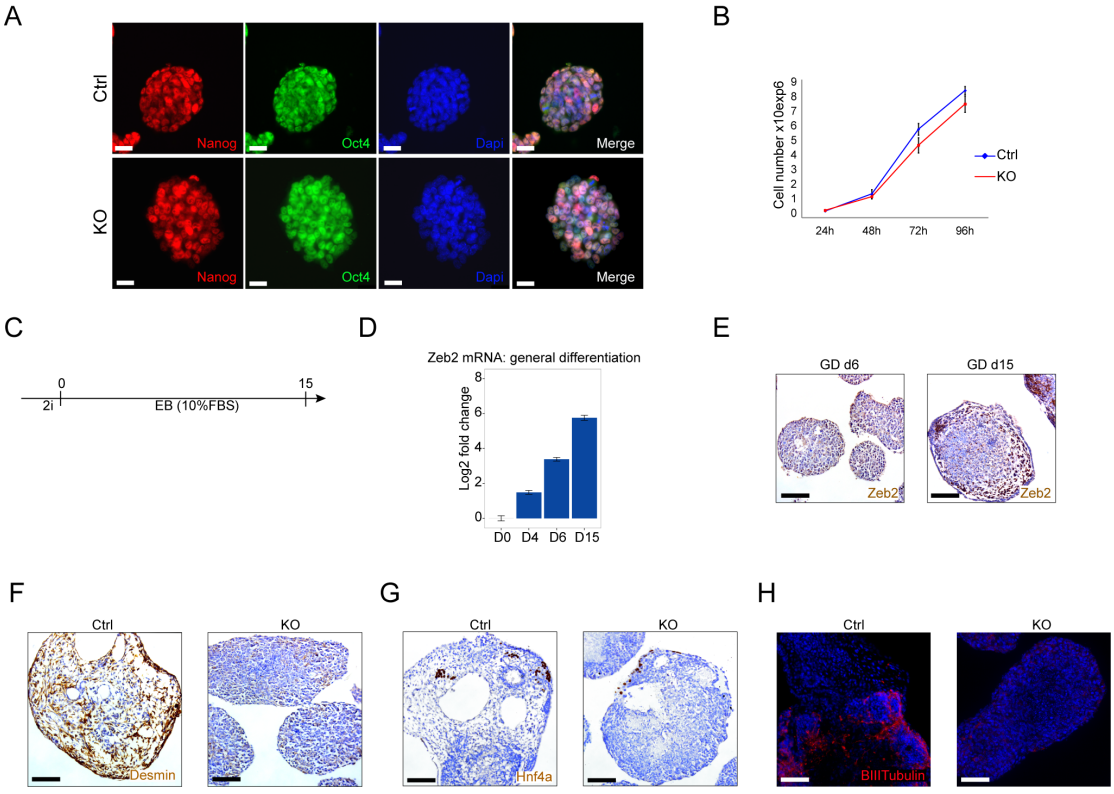

Figure 19. Genetic inactivation of *Zeb2* in mouse ESCs does neither impair ESC proliferation nor Oct4/Nanog production in pluripotency-supporting conditions.

A. Control (Ctrl) and *Zeb2* knockout (KO) ESCs, grown in LIF+2i, were co-stained by indirect IF for Nanog (red) and Oct4 (green). Results are representative of three experiments performed. Bar: 25 µm. B. Cell proliferation of control (blue) and *Zeb2* KO (red) ESCs grown in LIF+2i. SDs are of 3 biological replicates. C. Scheme of the general differentiation (GD) protocol. D. Zeb2 mRNA levels increased in Ctrl ESCs, albeit to lower relative levels as in ND (Fig. 1B). E. Ctrl EBs were stained for Zeb2 (brown) on d6 and d15 of GD. Scale bars: 50µm. F-H. IHC and IF analysis to document differentiation towards mesoderm (Desmin, panel F), endoderm (Hnf4α; panel G) and neural cells (BIII-Tubulin; panel H). Scale bars: 50µm. Results shown are from one experiment and are representative for 3 experiments. In contrast to Ctrl cells, very few Desmin+ cells were detected in *Zeb2* KO EBs after d15 in GD, while Hnf4α levels appear reduced in Zeb2 KO EBs. The neural commitment defects first observed in ND were re-confirmed in GD as BIII-Tubulin+ cells were absent in *Zeb2* KO EBs.

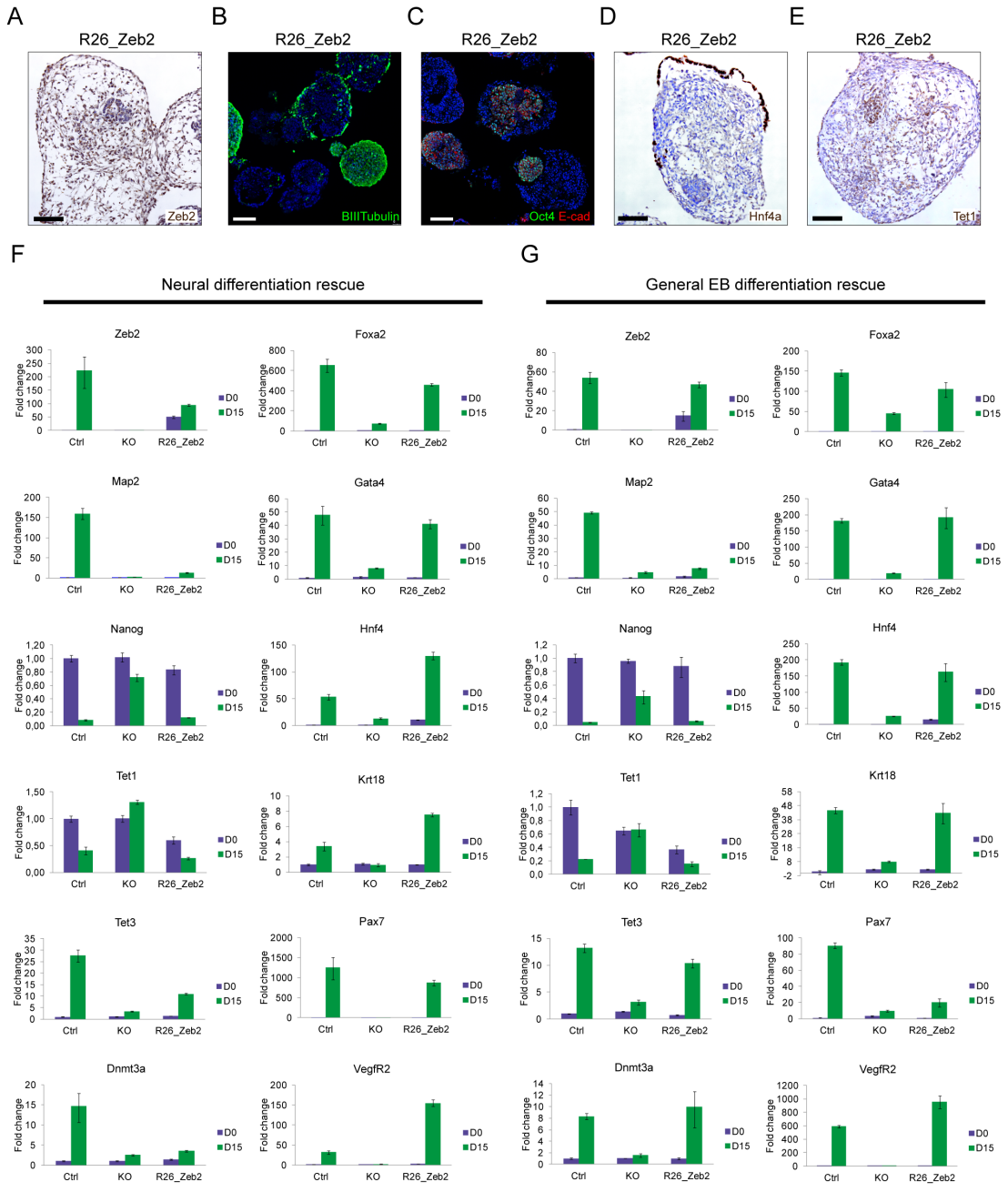

Figure 20. The differentiation defect is rescued in R26_Zeb2 ESC lines.

A-E. IHC and IF of R26_ZEB2 line after 15 days of general (GD) or neural (ND) differentiation A. Zeb2 (brown) in GD. B.IBIII-Tubulin (green), ND. C. Oct4 (green) and Cdh1 (red), ND. D. Hnf4a (brown), GD. E. Tet1 (brown). Scale bar: 50µm. F. RT-qPCR analysis of marker genes representative for three-lineage differentiation, pluripotency and DNA-methylation in Ctrl, *Zeb2* KO and R26_Zeb2 lines on d0 (violet) and 15 (green) of ND. Results are representative of three experiments performed. SD of two technical replicates is shown. G. RT-qPCR of marker genes representative for three-lineage differentiation, pluripotency and DNA-methylation in Ctrl, *Zeb2* KO and R26_Zeb2 lines on d0 (violet) and 15 (green) of GD. Results are representative of three experiments performed. SD of two technical replicates is shown.

4.3.2 Zeb2 acts preferentially as a transcriptional repressor associated with developmental progression

Temporal RNA-seq of both Ctrl and *Zeb2* KO ESCs showed in more detail Zeb2-dependent effects on early cell-state/fate decisions on the one hand and identify potential mediators of the impaired differentiation phenotype downstream of Zeb2 on the other hand. We chose ND where we can distinguish at least three stages that correspond in Ctrl cells to (i) ground-state ESCs (d0, very low Zeb2 mRNA/protein), (ii) multipotent progenitors (d4, low Zeb2) and (iii) early neural progenitors (d6, high Zeb2). For each stage we performed RNA-seq for three independent experiments. Principal component analysis (PCA) illustrated that both Ctrl and *Zeb2* KO on d0 are situated close together, but on d4 they already follow different lineage trajectories (Fig. 21A). This coincides with the first induction of *Zeb2* (between d0-d4 in Ctrl, see Fig. 18B) indicating that Zeb2 influences cell-fate decisions very early on when cells normally exit from their ground-state and undergo lineage priming.

We next applied a time-series analysis on our RNA-seq data set to identify genes that have a different dynamic expression profile in KO vs. Ctrl cells. Gene ontology (GO) analysis (using GSEA-P; Subramanian *et al.*, 2005) of genes displaying this different dynamic behavior showed strong negative enrichment for various differentiation/developmental categories within the top-10 hits (Fig. 21B) and further confirmed that at least the vast majority of *Zeb2* KO cells in EBs indeed remain uncommitted during differentiation. We also (re-)confirmed that *Zeb2* KO cells do not acquire neural fate using markers as *Pax6*, *Zfp521* and *Neurog1* (Fig. 22A), and the early-neuroectoderm markers *Gbx2* and *Hoxa1* previously shown to be correctly induced upon differentiation in Zeb2 KD hESCs (Chng *et al.*, 2010). *Gbx2* and *Hoxa1* were not induced in *Zeb2* KO mESCs, indicating that genetic inactivation of *Zeb2* results in a more severe neural acquisition phenotype than the KD (Fig. 22A). We examined the expression profiles of other cell lineage markers in our RNA-seq data to exclude that *Zeb2* KO cells would preferentially induce non-neural fates (Fig. 22B-D). Although a small increase in those markers was observed in Ctrl EBs, they were either almost absent (for mesoderm, Fig. 22C) or markedly lower (for trophectoderm and endoderm; Fig. 22B,D) in Zeb2-deficient EBs.

While this RNA-seq data analysis significantly expands our previous characterization of the cells and confirms that *Zeb2* inactivation globally affects ESC differentiation potential, it provided the possibility to discover potential Zeb2-dependent candidate genes responsible for the impaired differentiation of *Zeb2* KO ESCs. Therefore we performed pairwise RNA-seq analysis at all three time-points and identified differentially expressed genes (DEGs) (p-value < 0.01 and log2 Fold Change (FC) > 1) Consistent with the divergent PCA trajectories we observed an increase in both number of DEGs and their FC over time (Fig. 21C). Remarkably, on d0 when Zeb2 mRNA is almost undetectable and no clear cellular phenotype is observed, there are already 742 assignable DEGs, although more than 95%

of these DEGs display only small changes (<2 log2 FC). GO-analysis showed that almost all of these genes relate to metabolism or cell cycle regulation (*not shown*).

Upon neural induction the majority of genes that were either up or down on d4 (multipotent progenitor stage) maintained this status on d6 (early neural progenitor), 69% and 72%, respectively. Numbers of DEGs increased between d4 and d6. To further filter for direct transcriptional regulation by Zeb2 we performed a binding motif analysis within promoter regions (2kb up and downstream of the transcription start site, TSS) of DEGs. We searched for two motifs, the E-box sequence 5'-CACCTG-3' and the sequence 5'-CACCT-3', interspaced by maximum 45bp (Remacle *et al.*, 1999) The genes upregulated during differentiation in *Zeb2* KO cells showed an increase in enrichment for the selected Zeb2-binding motif (Fig. 21D, red bars and line), while the opposite trend was observed for the downregulated genes (Fig. 21D, green bars and line). Altogether this suggests that during differentiation Zeb2 functions preferentially as a transcriptional repressor.

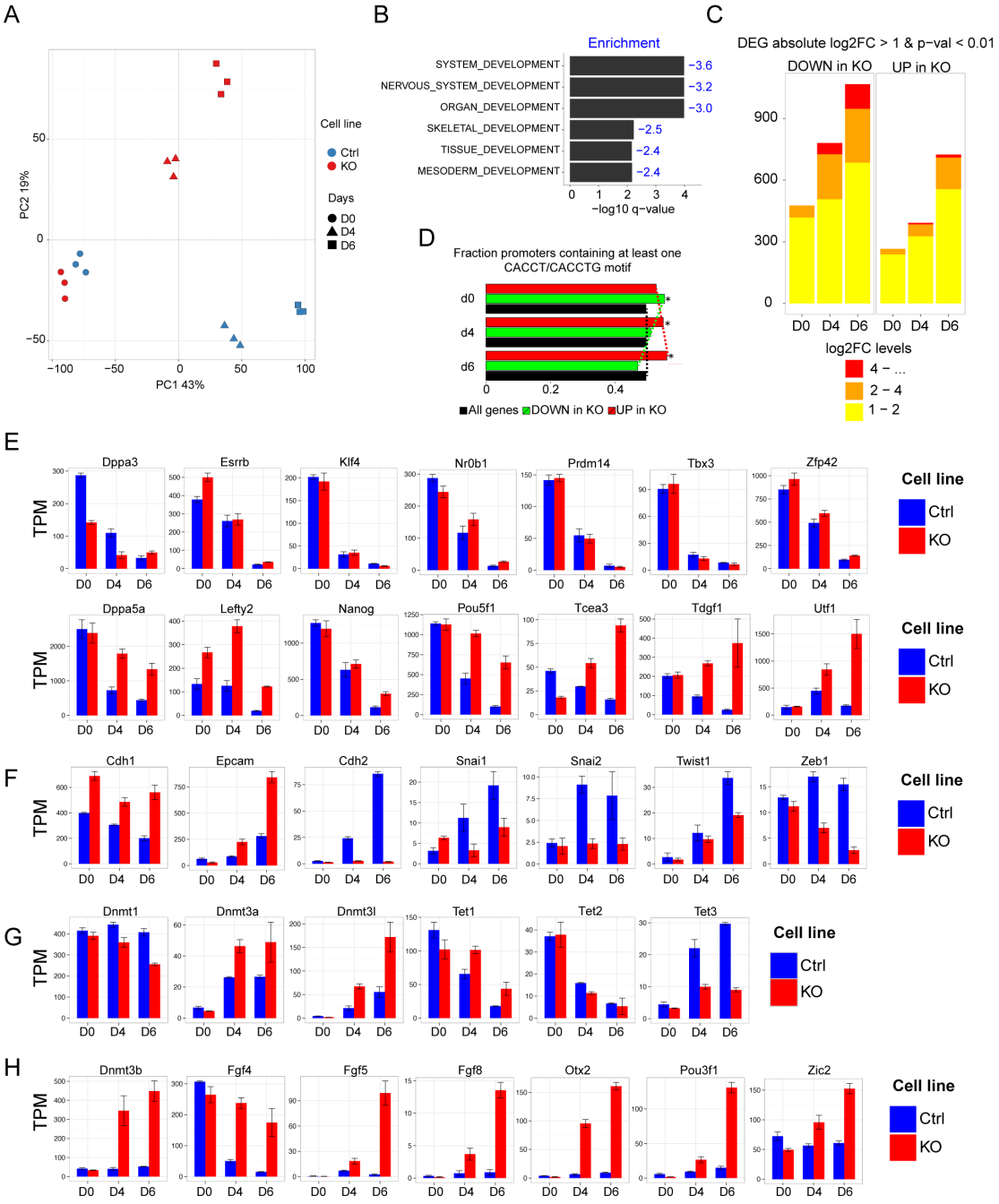

Figure 21. Analysis of temporal RNA-seq.

A. Principal component analysis based on transcripts per million (TPM). B. GSEA-P for DEG in time-series analysis. The height of the bar plot represents significance and the corresponding negative enrichment score is indicated (blue). C. Bar plot displays numbers of DEGs using pairwise DESeq2 test (|log2FC| > 1 and p < 0.01). Colors represent binned absolute log2FC levels. D. Promoter analysis for putative bipartite Zeb2-binding motifs (CACCT/CACCTG sequences with maximum gap of 45bp; see main text) of DEGs between *Zeb2* KO vs. Ctrl. Red bar = selective analysis for upregulated DEG, demonstrates statistical overrepresentation (Fisher's exact test p-value = 1.044e-08). Green bar = selective analysis for downregulated DEG points to underrepresentation. E-H. TPM (transcripts per million) bar plots at the indicated time points for pluripotency-related genes (E), selected EMT genes (F) and selected methylation-related genes (G) and epiblast (H).

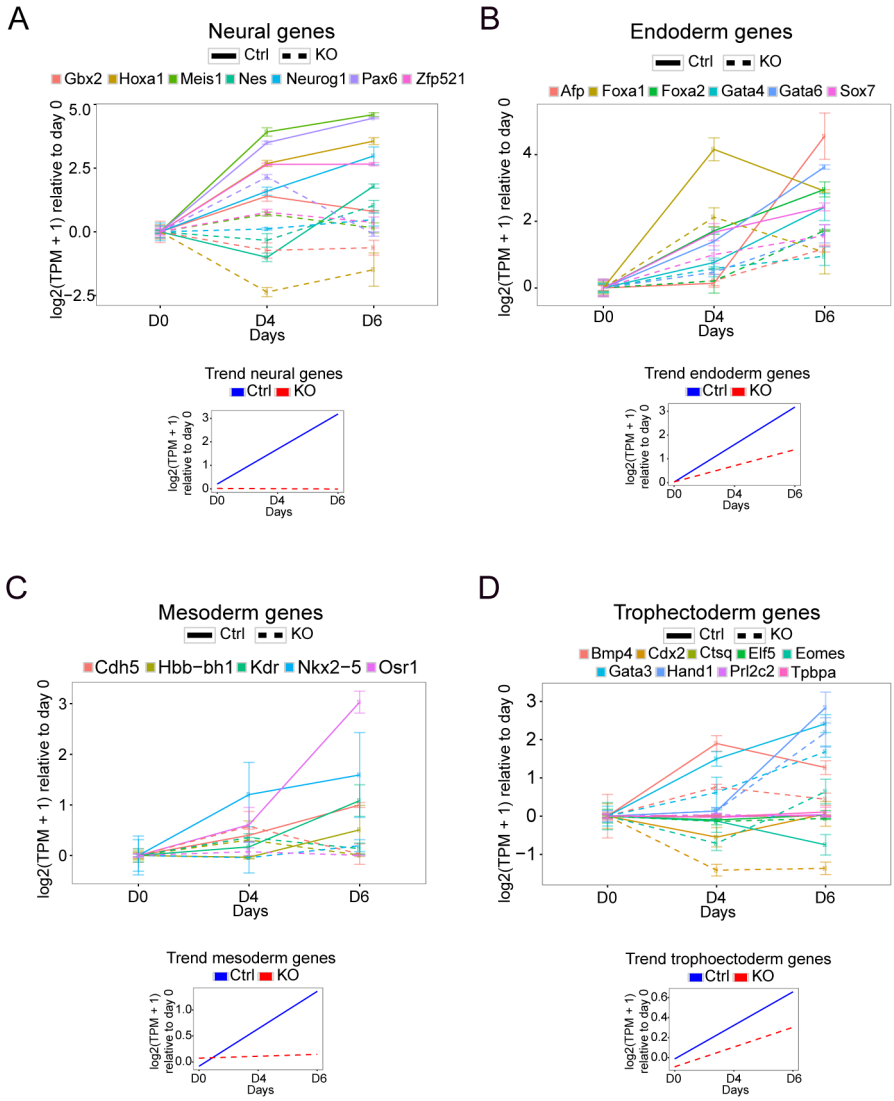

Figure 22. Dynamic expression levels in Ctrl and KO cells

for selected neural (A) endodermal (B), mesodermal (C) trophectodermal (D) genes on d0, d4, d6. Average trend for selected genes in Ctrl and KO cells is modeled by simple linear regression line.

4.3.3 *Zeb2* KO ESCs stall in an epiblast-like state

Zeb2 deficiency leads to impaired differentiation of ESCs and Zeb2 displays a preferential repressive function as shown above. We therefore investigated whether the pluripotency network was properly silenced in *Zeb2* KO ESCs, in particular the genes associated with the naïve state known to be rapidly downregulated upon withdrawal of LIF+2i (Leeb *et al.*, 2014). *Klf4*, *Tbx3*, *Zfp42*, *Prdm14*, *Essrb, Nr0b1* and *Dppa3* were all properly downregulated in both *Zeb2* KO and Ctrl ESCs (Fig. 21E, upper panel). However, a significant set of factors that are part of a larger pluripotency network or involved in initiation of differentiation were not at all or only partially downregulated, such as *Lefty2*, *Tcea3*, *Dppa5a*, *Utf1*

and *Tdgf1* (Chambery *et al.*, 2009; Jia *et al.*, 2012; Kim *et al.*, 2013; Park *et al.*, 2013; Qian *et al.*, 2015) (Fig. 21E,lower panel). This group also included *Pou5f1* and *Nanog*, key players in the acquisition of pluripotency and early development (Chambers *et al.*, 2003; Nichols *et al.*, 1998). All genes in the latter group contain putative binding sites for Zeb2 within 2kb from their transcription start sites (TSS), suggesting that Zeb2 is a candidate direct repressor of (at least some) genes involved in pluripotency maintenance. In line with the known role of Zeb2 in EMT (Comijn *et al.*, 2001; Lamouille *et al.*, 2014) we observed that in *Zeb2* KO cells expression of *Cdh1* remains high, *Epcam* is strongly induced and *Cdh2*, *Snai-1*/*2*, *Twist1* and *Zeb1* were not induced to the same extent in differentiation conditions (Fig. 21F). Altogether these observations confirm that these cells have defective EMT consistent with previously documented roles of Zeb2, including downregulation of *Cdh1*, in other cell types.

Both *Dnmt3b* (Fig. 21H) and *Dmt3l* (Fig. 21G) have putative Zeb2-binding sites and were upregulated in *Zeb2* KO during differentiation. Together with other genes they are important for DNA-methylation at this stage, hence we also monitored the expression of *Dnmt1*, *Dnmt3a* and *Tet1*, *Tet2* and *Tet3* (Okano *et al.*, 1999; Jackson *et al.*, 2004; Tahiliani *et al.*, 2009; Ito *et al.*, 2010; Koh *et al.*, 2011). In addition to the high expression of all three *Dnmt3* genes, the *Tet1/2* to *Tet3* expression switch is only partial. It normally occurs during transition from pluripotent stem cells to differentiated cells (Koh *et al.*, 2011), but in our case *Tet3* induction is limited and *Tet1* expression is higher (Fig. 21G).

Although EBs are inherently heterogeneous, all our previous data indicates that at least part of the cells in EBs are stalled in an epiblast-like cell state *in vitro,* but which cannot be propagated as long-term epiblast-like cells (see *e.g.*. Hayashi *et al.*, 2012 and references therein). This is further supported by a strong increase of expression of the established post-implantation epiblast genes *Otx2*, *Pou3f1 (Oct6)*, *Dnmt3b, Zic2* and *Fgf5* (Buecker *et al.*, 2014; Iwafuchi-Doi *et al.*, 2012) (Fig. 21H). Interestingly, we noticed that in addition to *Fgf5* the epiblast-linked genes *Fgf8* and *Fgf4* (Joo *et al.*, 2014) also show increased mRNA levels with putative Zeb2-binding sites in the promoter, but whether the maintenance of these deviating FGF signals help to maintain the epiblast-like cell state of the Zeb2-deficient ESCs remains to be investigated.

4.3.4 Pluripotency genes are not repressed during differentiation of Zeb2KO ESCs

Both Pou5f1 (Oct4) and Nanog, two crucial pluripotency-supporting factors maintained in epiblast cells, remained high as seen by western blotting and RT-qPCR (Fig. 24A,B) and were present in a large fraction of cells in *Zeb2* KO ND-EBs till d15 (Fig. 23A,B). Again this observation could be extended to GD-EBs (Fig. 24C,D, shown on d15). In addition high numbers of Cdh1+ cells were observed in *Zeb2*

KO EBs (Fig. 23B) and this was also seen at protein and mRNA levels on d15 (Fig. 24A,B). R26_Zeb2 rescue restored downregulation of Oct4 (and Cdh1) mRNA/protein (Fig. 20C). To confirm that this is the direct result of Zeb2-binding a ChIP-qPCR was carried out over the Zeb2-binding motif (Fig. 23C,D). This showed enrichment of Flag-tagged Zeb2 (using R26_Zeb2 ESCs) in the promoter of *Cdh1* (Comijn *et al.*, 2001; Van de Putte *et al.*, 2003) and its new candidate target *Nanog* in ESCs.

To document the persistence of the pluripotent state upon differentiation in *Zeb2* KO cells, we dissociated Ctrl and *Zeb2* KO EBs on d15 (either in ND or GD), sorted the living cells and plated these at 500 cells/well as single cells in 2i. Alkaline phosphatase (AP)+ ESC colonies derived from EBs subjected to differentiation (Fig. 23E,F) were quantified on d9 (Fig. 23G). In a typical experiment, Ctrl cells subjected to ND did not give rise to AP+ cells, whereas *Zeb2* KO cells in 2i yielded on average 8 colonies/well. In GD, Ctrl cells gave rise to less than 1 (calculated 0.2) AP+ colony/well, whereas for *Zeb2* KO cells this was 4 colonies/well on average. Based on AP read-out, this shows they have the remarkable ability to re-adapt to 2i, like ESCs and EpiSCs and form AP+ colonies even up to d15 of differentiation treatment. Without assessment at single-cell level, we cannot discriminate whether these AP+ colonies arose exclusively from epiblast-like or more naïve *Zeb2* KO cells since both cell types can adapt to 2i. Teratoma formation, using EBs subjected to ND for 12 days showed that Ctrl EBs failed to form teratomas, while *Zeb2* KO EBs gave rise to teratomas in 4 weeks (Fig. 24E).

These data show therefore that Zeb2 deletion leads to maintenance of pluripotency even after prolonged *in vitro* differentiation.

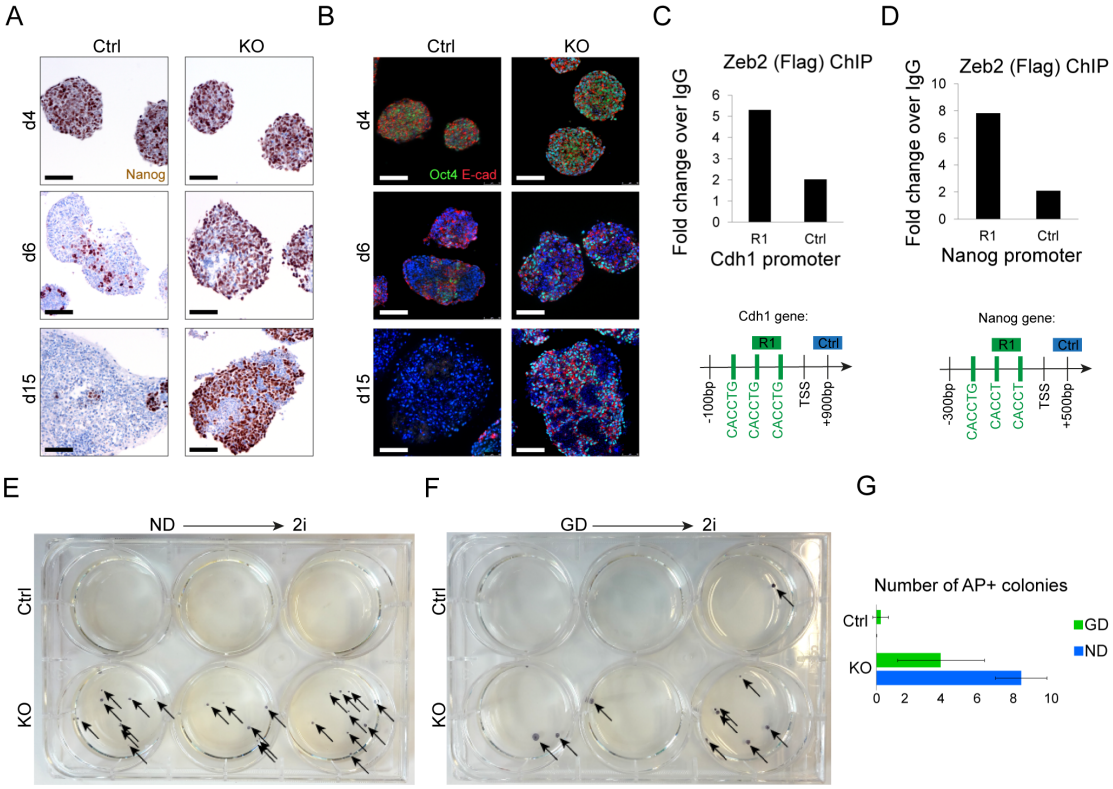

Figure 23. Pluripotency genes are not repressed during differentiation in the Zeb2 KO ESCs.

Pluripotency genes are not efficiently downregulated during differentiation in *Zeb2* KO ESCs. A. Control (Ctrl) and *Zeb2* KO (KO) EBs stained for Nanog (brown) on d4, d6 and d15 of ND. B. Ctrl and KO EBs co-stained for Oct4 (green) and Cdh1 (red) on d4, 6 and 15 of ND. Panels A-B show results from one experiment that is representative for 3 experiments. Scale bar: 50µm. C-D. Zeb2 ChIP (using anti-Flag antibody) on *Cdh1* (panel C) and *Nanog* promoter (panel D). Results shown are from one experiment and representative for 3 experiments. E-F. Ctrl and KO ESCs subjected to ND (panel E) and GD (panel F) for 15 days, dissociated and plated at 500 cells/well in 2i. The resulting ESC colonies (indicated by arrows) were visualized by staining for AP, and panel G represents the average number of AP+ colonies obtained after plating the cells.

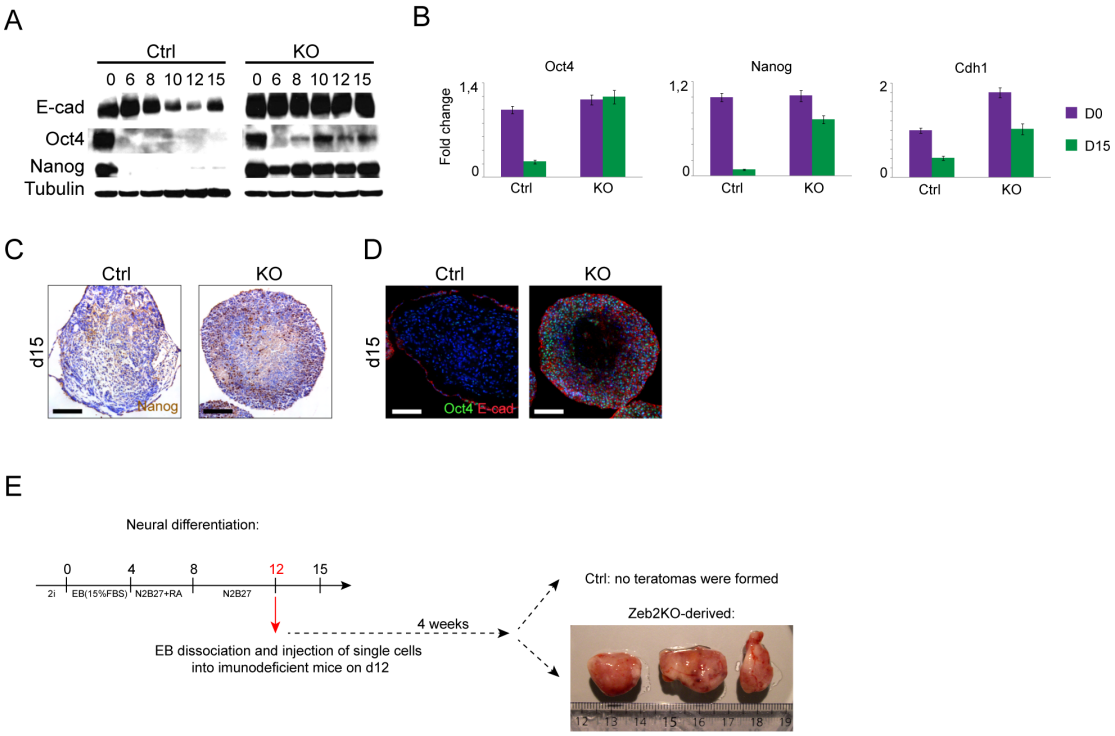

Figure 24. Pluripotency gene expression is not silenced during differentiation in *Zeb2* knockout (KO) cells.

A. Western blot analysis for Cdh1, Oct4 and Nanog in Ctrl and *Zeb2* KO (KO) EBs during ND. B. RT-qPCR analysis of Ctrl and KO ESCs (D0) and EBs on d15 of ND for Oct4, Nanog and Cdh1. Results are representative of three experiments performed. SD of two technical replicates is shown C. Ctrl and *Zeb2* KO EBs stained for Nanog (brown) on d15 of GD D. Ctrl and KO EBs co-stained for Oct4 (green) and Cdh1 (red) on d15 of GD. Panels C-D show results from one experiment that is representative for 3 experiments. Scale bar: 50µm E. Teratoma formation assay with EBs subjected to ND for 12 days.

4.3.5 The *Zeb2* KO embryoid bodies, subjected to neural differentiation, fail to maintain the initially acquired DNA-methylation

Zeb2 KO EBs show deregulated expression of the core methylation machinery genes together with impaired differentiation. This prompted us to examine the acquisition and maintenance of CpG-methylation (^{me}CpG) that accompanies irreversible ESC differentiation decision. Retaining the same time/sample setups as for RNA-seq and again using ND, single-base profiles were generated of methylation by RRBS in both Ctrl and *Zeb2* KO on d0, d4 and d6.

The genome of ground-state (d0) ESCs was globally hypomethylated (see also Hackett *et al.*, 2013). On d4 both cell populations gained methylation in agreement with our observation that they are in an epiblast-like (for KO) or multipotent (Ctrl) state. However a significant drop of ^{me}CpG was observed in the d6 *Zeb2* KO cell population, suggesting that part of the CpG methylation is lost again (Fig. 25A). The progressive accumulation of ^{me}CpG in our Ctrl EBs has a striking resemblance with that observed

in vivo (Auclair *et al.*, 2014), i.e. our d0 population profile is similar to blastocyst-stage embryos (between E3.5-E4.5), while d4 and d6 Ctrl populations have a similar distribution profile as epiblast embryos (E6.5). In contrast KO d6 resembles early-epiblast embryos (E5.5) with a reduction in ^{me}CpG at both gene bodies and 10kb-flanking regions (Fig. 25B) (Auclair *et al.*, 2014).

We further investigated changes in the (de)methylation process by considering CpGs covered in all samples and averaging methylation in 400bp-tiles, with a total of 184564 tiles. This identified differentially methylated regions (DMRs) (absolute methylation change >20% and q-value <0.05) in both a time and pairwise-dependent manner. Both Ctrl and KO cells significantly gained methylation in respectively 33.8% and 33.5% of all tiles between d0 and d4 (Fig. 25C, left panels and Fig. 26A). During this period no single significant loss-of-methylation was observed (Fig. 26B). Next, between d4 and d6 Ctrl cells maintained a very stable level of methylation with only little gain or loss-of-methylation, i.e. 0.1% of all tiles (Fig. 25C, right top panel and Fig. 26A-B). In agreement with the observed overall lower methylation at d6 (see Fig. 25A-B), KO cells had 10 times more tiles (1806 or 1% of all tiles) with significant loss-of-methylation and only 90 tiles (0.05% of all tiles) with gain-of-methylation (Fig. 25C, right bottom panel and Fig. 26A-B,). Furthermore, analysis of these aforementioned DMR in both Ctrl and KO cells revealed that these regions initially acquired methylation in both Ctrl and KO cells at d4, but this methylation was only maintained in the Ctrl cells (Fig. 26C).

To investigate whether demethylation was selective for specific genomic regions, we profiled the methylation dynamics of enhancers, CpG islands (CGI), canyons, transposable elements (TE), high CpG-content (HCP) and low CpG-content promoters (LCP), exons and introns. As reported before (Jeong *et al.*, 2014b) resistance to gain-of-methylation occurs for canyons and high-CpG regions (CGI and HCP), while all other regions (enhancers, TE, LCP, exons and introns) were susceptible to gain-of-methylation. In contrast, the *Zeb2* KO population is unable to maintain this methylation initially acquired in all aforementioned genomic regions (Fig. 25F).

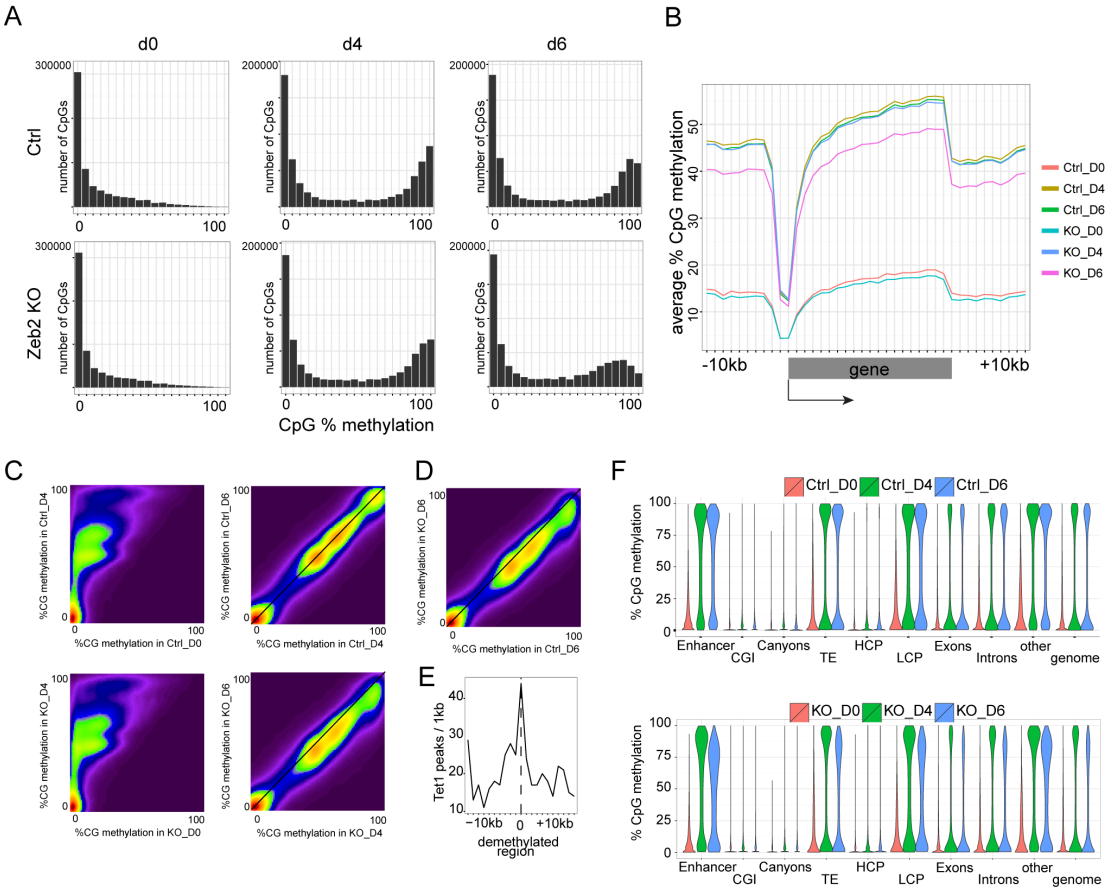

Figure 25. Analysis or temporal RRBS during neural differentiation.

A. Distribution histogram for individual ^{me}CpGs on d0, d4 and d6 in Ctrl and *Zeb2* KO populations. B. ^{me}CpG distribution at gene bodies and 10kb-flanking regions of protein-coding genes. C. Density plots for pairwise comparisons of ^{me}CpG (in 400bp-tiles) between d0-d4 in Ctrl (top) and d4-d6 in KO (bottom) cells. D. Density plot for pairwise comparison of ^{me}CpG (in 400bp-tiles) on d6 between Ctrl and KO. In C,D the density points increase from purple to dark red. E. Enrichment plot of Tet1-binding peaks centered around demethylated regions on d6 in a pairwise comparison between *Zeb2* KO vs. Ctrl. F. Violin plots showing gain and loss-of-methylation over time in identified genomic regions, *i.e.* enhancers, CpG islands (CGI), canyons, transposable elements (TE), high-CpG content promoters (HCP), low-CpG content promoter (LCP), exons, introns, other non-defined genomic regions, and globally at the whole-genome (genome) in Ctrl and *Zeb2* KO cells.

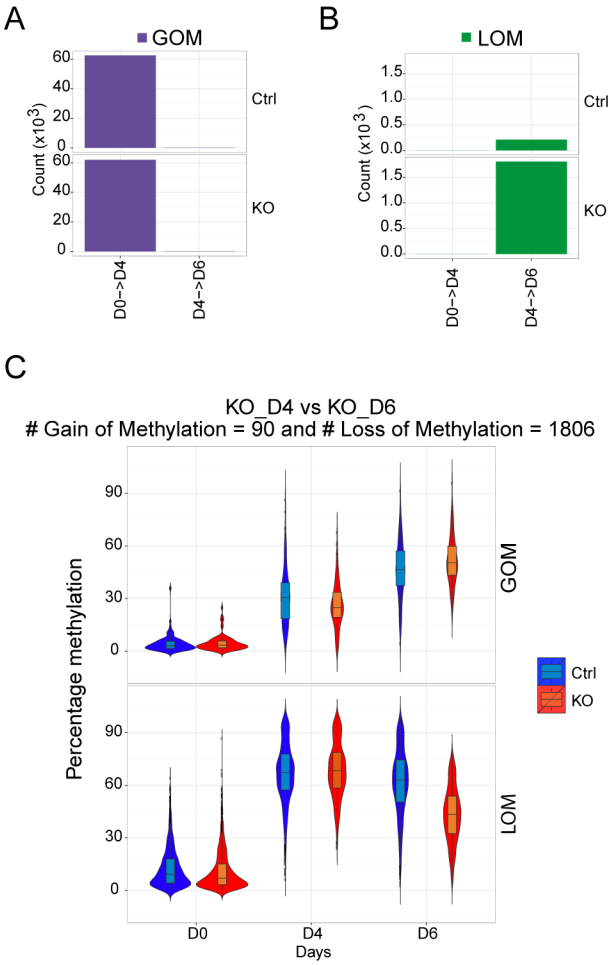

Figure 26. RRBS.

A-B. Bar plot showing gain (GOM, panel A) and loss-of-methylation (LOM; panel B) between consecutive time points in Ctrl and *Zeb2* KO ESCs. C. Violin plots to illustrate distribution and dynamic behavior of DMRs (both GOM and LOM) over time.

4.3.6 Failure to maintain acquired DNA-methylation during neural differentiation is associated with Tet1-binding

We have shown that regions that lose methylation in d6 *Zeb2* KO populations initially indeed acquired methylation comparable to Ctrl (see Fig. 26C). Secondly, we also compared d6 of both Ctrl and *Zeb2* KO populations and as expected observed a similar number of tiles with loss-of-methylation (1938, or 1% of all tiles) (Fig. 25D). We observed also an increased level of Tet1 (see Fig. 21) and therefore asked whether the regions that lose methylation correlate with Tet1-binding. Fig. 25E shows that regions that lose methylation are enriched for Tet1-binding by combining analysis of a published ChIP-seq data for Tet1 in mESCs (Wu *et al.*, 2011) with our region-specific loss of methylation data on d6 (compared between Ctrl and *Zeb2* KO). This strongly suggests that the observed demethylation in the

Zeb2 KO cells is an active process mediated by elevated Tet1 levels in agreement with previous observations that DNA-demethylation is initiated at Tet1-binding sites (Habibi *et al.*, 2013).

4.3.7 Tet1 knockdown in *Zeb2* KO ESCs facilitates silencing of *Nanog*, *Oct4* and *Cdh1* and partially rescues the lineage differentiation phenotypes

Tet1 remains high in the *Zeb2* KO EBs even on d15 of differentiation in contrast to the normal downregulation during ND and GD (Fig. 27A,B). To test whether the high Tet1 levels lead to inefficient silencing of *Nanog*, *Oct4* and *Cdh1* and hence a block in the differentiation potential of these cells, we transduced control and *Zeb2* KO ESC lines with a lentivirus expressing shRNA directed against Tet1 (hereafter called Ctrl_Tet1shRNA, Zeb2KO_Tet1shRNA lines, respectively). Tet1 protein was almost undetectable in the Ctrl and *Zeb2* KO lines targeted with this Tet1 shRNA (see Fig. 27C and Fig. 27D). In 2i conditions, the Tet1 KD lines maintained their undifferentiated characteristics (*not shown*).

We subjected these Tet1shRNA lines to ND and GD, respectively, along with the same lines with a control non-targeting shRNA (Ctrl_CtrlshRNA and Zeb2KO_CtrlshRNA) These control shRNA lines behaved as expected in differentiation (Fig. 27E-J), and Zeb2 was indeed absent from Zeb2KO_CtrlshRNA and Zeb2KO_Tet1shRNA EBs at the end of GD (Fig. 27H) and ND (*data not shown*). After 15 days Zeb2KO_Tet1shRNA cells subjected to either ND or GD efficiently decreased Nanog, Oct4 and Cdh1 mRNA to low levels at the end of GD (Fig. 27E, F; ND, *data not shown*). In Zeb2KO_Tet1shRNA lines subjected to GD, partial rescue of differentiation to mesoderm (Fig. 27I) and endoderm (Fig. 27G) was observed, but not to neuroectoderm, (*data not shown*). Partial rescue of ND was observed only when Zeb2KO_Tet1shRNA cells were subjected to ND (Fig. 27J). Thus, Tet1 remains high in *Zeb2* KO cells during differentiation, and forced downregulation of Tet1 in these cells in such conditions enables the decrease of *Nanog*, *Oct4* and *Cdh1* transcription and partially rescues cell differentiation. We conclude that deficiency in Zeb2 during differentiation leads to higher Tet1 levels, which are associated with improper reduction of *a.o.* key pluripotency genes *Nanog* and *Oct4,* resulting in impaired differentiation.

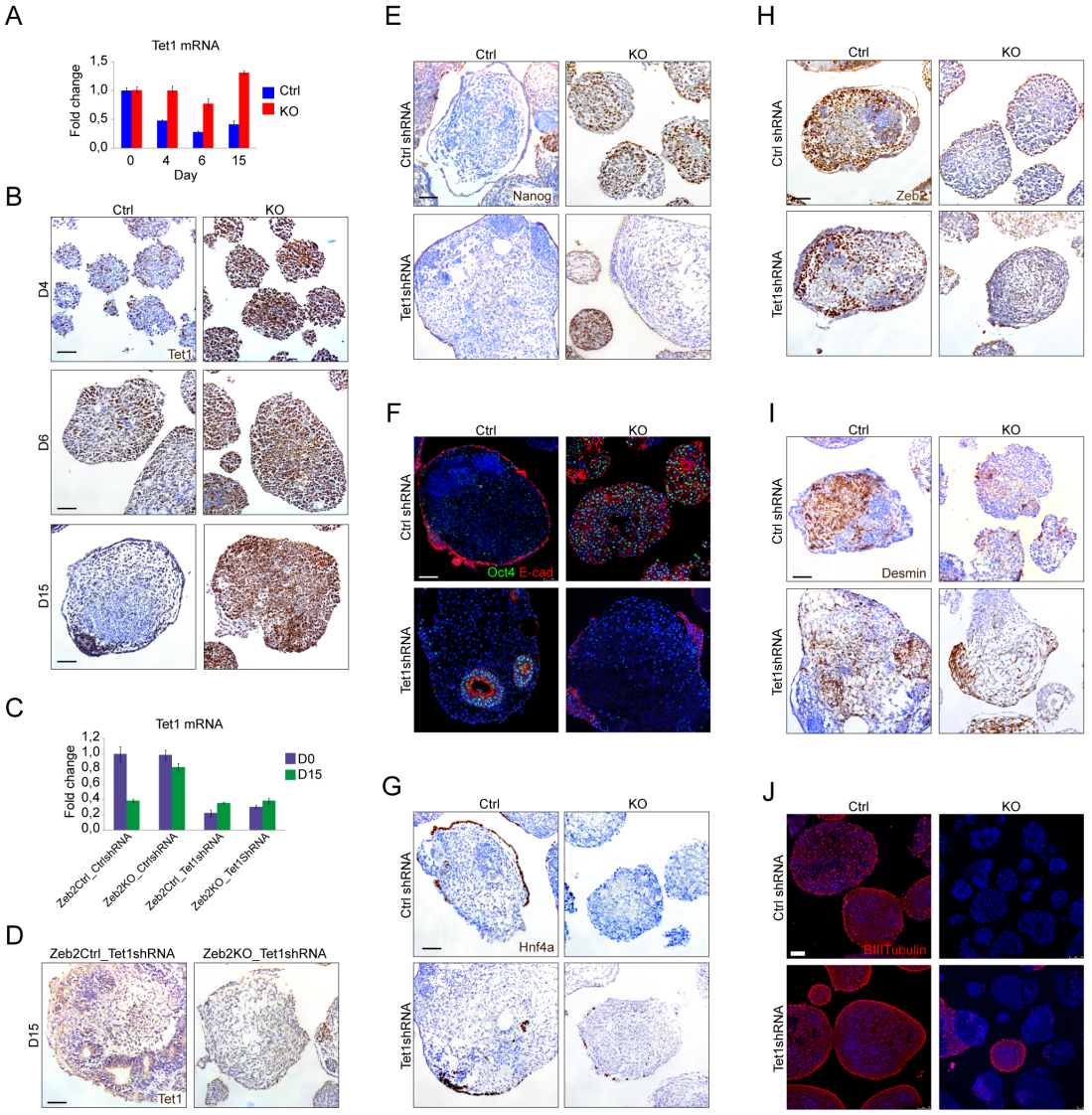

Figure 27. Tet1 knockdown in the Zeb2 KO cells facilitates exit from pluripotency and partially rescues their differentiation capacity.

A. RT-qPCR of Tet1 mRNA in Ctrl (blue) and *Zeb2* KO (red) lines on d0, d4, d6 and d15 of ND. SD of two technical replicates is shown. B. Ctrl and *Zeb2* KO EBs stained for Tet1 (brown) on d4, d6 and d15 of GD. C. RT-qPCR of Tet1 mRNA on d0 (violet) and d15 (green) in Ctrl_CtrlshRNA, Zeb2KO_CtrlshRNA, Ctrl_Tet1shRNA and Zeb2KO_Tet1shRNA lines. SD of 2 technical replicates is shown. D. Ctrl_Tet1shRNA and Zeb2KO_Tet1shRNA EBs co-stained for Tet1 (brown) on d15 of GD. Scale bar B,D: 50µm. E-J: Ctrl_CtrlshRNA, Zeb2KO_CtrlshRNA, Ctrl_Tet1shRNA and Zeb2KO_Tet1shRNA EBs stained for the indicated markers. E. Nanog (brown) on 15 of GD. F. Oct4 (green) and Cdh1 (red) on d15 of GD. G. Hnf4a (brown) on d15 of GD. H. Zeb2 (brown) on d15 of GD. I. Desmin (brown) on d15 of GD. J.BIII-Tubulin (red) on d15 of ND. Scale bar: 75µm. E-I: scale bar: 50µm. In all panels, results shown are from one experiment and representative for 3 experiments.

4.3.8 The neuronal inhibitory gene *REST* is deregulated in *Zeb2* KO ESCs

Repressor element 1 (RE-1) silencing transcription factor (REST) (Chong *et al.*, 1995), also known as neuron-restrictive silencer factor (NRSF; Schoenherr and Anderson, 1995) is, as these names suggest, a negative regulator of the expression of many genes in non-neural cells during neurogenesis, and the *Rest* KO mouse has bot defects in non-neural as well as neural cells (Chen *et al.*, 1998). REST binds to a 21 bp-long sequence, referred to as RE-1 (TCCAGCACCACGGACAGTTCC), in its target loci.

My RNA-seq analysis showed that *Rest* was moderately, but detectably upregulated in the *Zeb2* KO EBs during ND (d4 FC=1.70; d6 FC=1.77). I therefore documented Rest mRNA expression changes in control and *Zeb2* KO cells during 15 days of differentiation. Rest mRNA levels were comparable in the two lines during the first 3 days, however from d4 these mRNA levels decreased in the control cells, but not in the *Zeb2* KO cells where it was maintained at relatively high levels throughout. Western blot analysis showed that Rest was abundantly present in these ESCs. The levels decreased as the control ESCs differentiated and were undetectable from d8. *Zeb2* KO cells thus maintained high Rest levels during 15 days of ND. We reasoned that the defect in ND in the *Zeb2* KO cells would in part be due to the lack of repression activity by Rest.

4.3.8.1 Rest knockdown in the Zeb2 KO ESCs partially rescues neural differentiation

We generated REST knockdown lines by transducing control and *Zeb2* KO ESCs, respectively, with a lentivirus expressing short hairpin RNA (shRNA) directed against Rest (hereafter called Ctrl_RESTKD, Zeb2KO_RESTKD lines, respectively). Rest mRNA levels were found to reduce with this shRNA to about 80% of the normal level in the Rest KD lines (Fig. 28A,C).

Next, we differentiated these Rest KD lines along with the same lines that received non-targeting (negative control) shRNA (Ctrl_CtrlKD and Zeb2KO_CtrlKD) (also used in the Tet1 KD experiment; see section 4.3.7). We observed partial rescue of neural differentiation (Fig. 28D) in the Zeb2KO_RESTKD line, however pluripotency markers (Oct4 and E-cad) were not significantly downregulated in this line (Fig. 28E). This indicated that the expression of these pluripotency-associated genes is not Rest-dependent. Furthermore, Rest is likely to play a role in the *Zeb2* KO ESCs that are primed and hence more likely to enter the differentiation program as compared to the ground-state ESCs where the pluripotency network is stable.

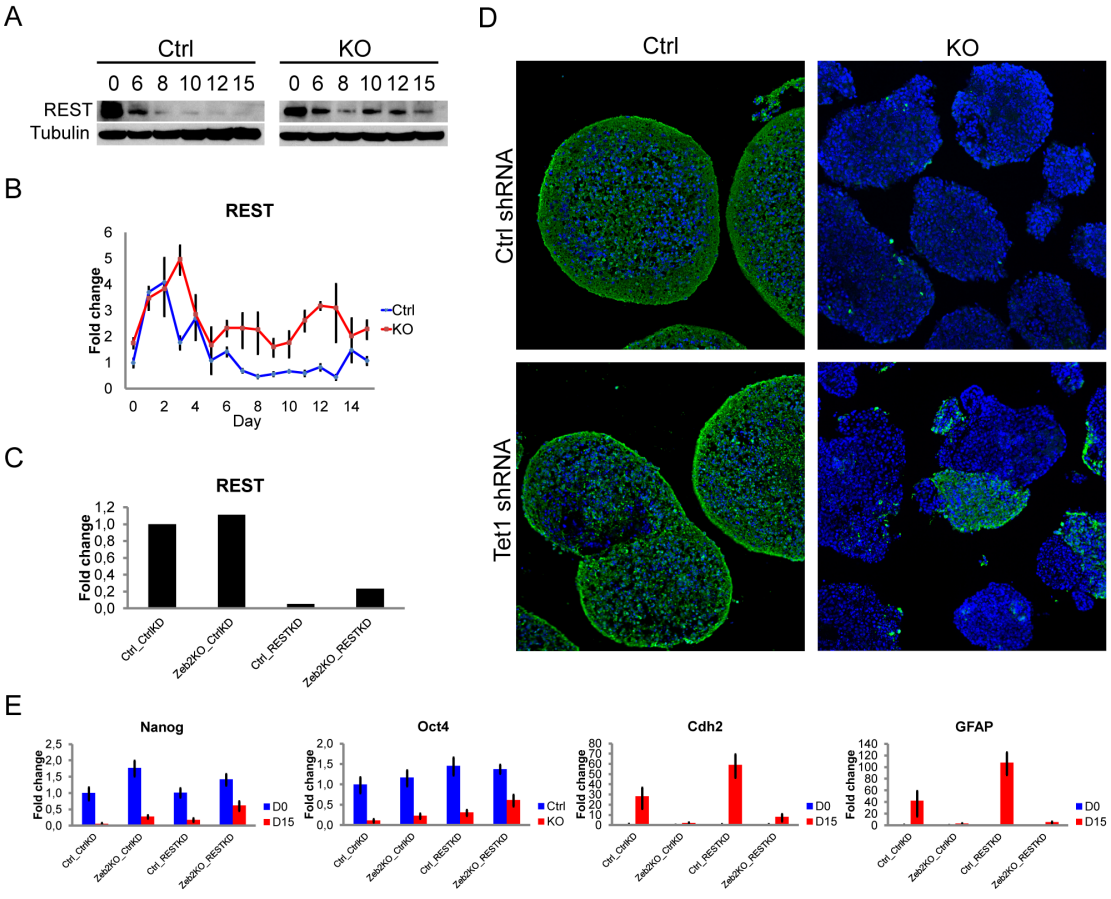

Figure 28. Rest knockdown in the *Zeb2* KO ESCs partially rescues their neural differentiation capacity.

A. Western blot analysis of Rest protein levels during ND of control and Zeb2 KO ESCs. B. RT-qPCR analysis of Rest mRNA expression during ND in control (in blue) and Zeb2 KO cells (in red). C. Rest was efficiently silenced in Ctrl (Ctrl_RESTKD) and Zeb2 KO (Zeb2KO_RESTKD) ESCs D. Ctrl_CtrlKD, Zeb2KO_CtrlKD, Ctrl_RESTKD and Zeb2KO_RESTKD embryoid bodies stained for BIII-Tubulin (green) on 15 of neural differentiation. Results are representative of three experiments. E. RT-qPCR of Nanog, Oct4, Cdh2, GFAP expression in Ctrl_CtrlKD, Zeb2KO_CtrlKD, Ctrl_RESTKD and Zeb2KO_RESTKD cells on d0 (in blue) and d15 (in red) of ND.

4.3.9 *Zeb2* KO ESCs have the capacity to differentiate into three lineages when injected into immunodeficient mice

Teratoma formation assays are used to demonstrate multi-lineage differentiation potential of ESCs. We subcutaneously injected 10^6 ESCs into immune-deficient mice (two ESC lines per animal, three animals per experiment; obviously one body side injected with one line). Teratomas formed in a period of 5 weeks. We then examined the gross histology of the control and *Zeb2* KO-derived teratomas, respectively. Both these lines gave rise to derivatives of mesoderm, endoderm and ectoderm (Fig. 29A). This indicated that *Zeb2* KO ESCs-when exposed to the specific differentiation queues provided by the host animal - regain their neural differentiation capacity, which is lost *in vitro*. We examined the presence of Zeb2 protein in the control and *Zeb2* KO-derived teratomas.

Control teratomas - as expected - showed high levels of Zeb2 in various tissues, including in cells of ectodermal origin. Surprisingly, *Zeb2* KO-derived teratomas also contained Zeb2+ cells indicating that the host cells mixed with injected ESCs (Fig. 29B). To exclude that the neural tissue present in the *Zeb2* KO teratomas was derived from the de-differentiated host cells and show that ESCs lacking Zeb2 had the capacity to generate neural tissues, we stably transduced *Zeb2* KO ESCs with a lentivrus encoding GFP (hereafter called Zeb2KO-GFP) and repeated the teratoma formation assays, and monitored the GFP+ teratomas. Neural tissue in the Zeb2KO-GFP teratomas was positive for GFP indicating that indeed, the *Zeb2* KO cells underwent neurogenesis *in vivo* (Fig. 29C). We further show - by IF analysis for BIII-Tubulin (neurons), Olig2 (oligodendrocytes) and GFAP (astrocytes) - that the *Zeb2* KO derived GFP+ ESCs generate all three neural cell lineages (Fig. 29D).

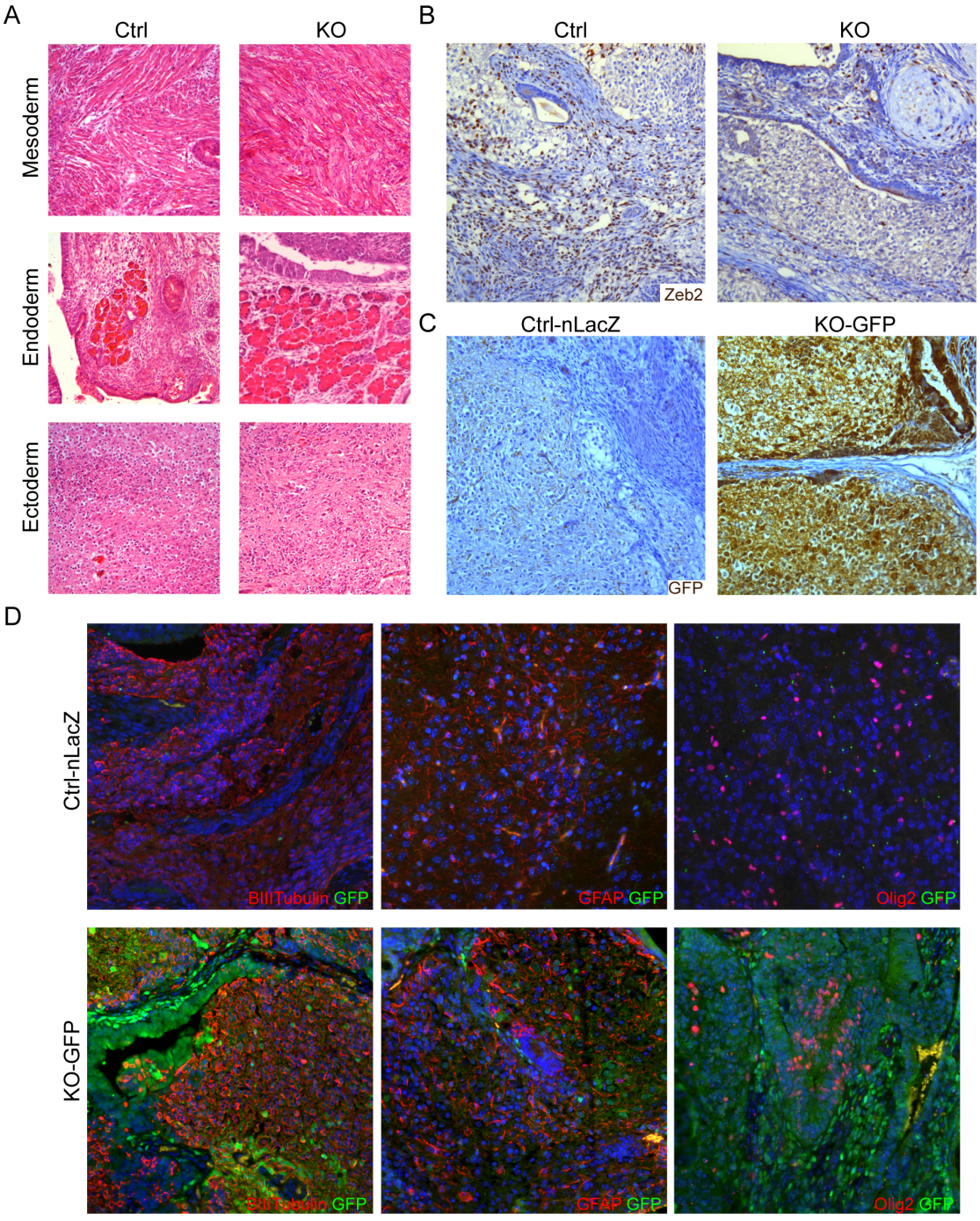

Figure 29. *Zeb2* KO ESCS differentiate to endoderm, mesoderm and ectoderm when injected into immunodeficient mice.

A. H&E analysis of the Ctrl and Zeb2 KO-derived teratomas. Three germ lines are present. B. Zeb2 staining (brown) in Ctrl and Zeb2 KO-derived teratomas C. GFP staining (brown) of the neural tissue in Ctrl-nLacZ and Zeb2KO-GFP derived teratomas. D. IF analysis of BIII-Tubulin (red, left column), GFAP (red, middle column), Olig2 (red, right column) and GFP (every column) of Ctrl-nLacZ and Zeb2KO-GFP-derived teratomas.

4.3.10 Zeb2 action *in vitro* is primarly cell-autonomous

Zeb2 was shown to act both in cell autonomous (van den Berghe *et al.*, 2013) and in cell non-autonomous fashion (Seuntjens *et al.*, 2009) during forebrain development. *Zeb2* KO ESCs show defects in *in vitro* neural and astroglial cell differentiation (see Chapter 3 and Introduction), however, the same cells - when injected into immunodeficient mice - are capable of generating neurons, astrocytes and oligodendrocytes within the developing teratomas (see Fig. 29), which indicates that the micro-environment provided by the recipient mouse is sufficient to repress the pluripotency network in the *Zeb2* KO ESCs and induce and efficiently execute the three-lineage differentiation program.

In order to recreate the wild-type environment that would provide *Zeb2* KO ESCs with sufficient amounts of growth factors that would lead to ESC-to-neuron cell fate switch *in vitro*, we performed co-culture experiment where control ESCs were mixed with *Zeb2* KO ESCs in different ratios at the onset of ND. To distinguish between the two lines, we provided a stable marking of control ESCs with nuclear LacZ (Ctrl-nLacZ; Fig. 30D) and *Zeb2* KO ESCs with GFP (Zeb2KO-GFP; Fig. 30E), respectively, using lentiviral transfer prior to the co-culture experiment. ESCs were mixed in the following % ratios (Ctrl/KO): 100/0, 90/10, 50/50,10/90 and 0/100, respectively. The Ctrl/KO aggregates formed normally and were morphologically indistinguishable from the EBs formed by either Ctrl-nLacZ or Zeb2KO-GFP ESCs. Zeb2KO-GFP cells were randomly distributed throughout the EBs in the first 2-3 days of differentiation. However, on d4 *Zeb2* KO cells in the EBs begun to cluster together (See Fig. 30A). Control cells migrated to the outside of the aggregates forming their outermost layer, whereas *Zeb2* KO cells formed tight clusters of GFP+ cells located more centrally within the EBs (Fig. 30 A,B,C,F,G). Initial observations suggested that Zeb2 regulated neurogenesis in a cell-autonomous fashion (Fig. 30H,I) since no GFP+ neurons (derived from Zeb2KO-GFP ESCs) were present in the aggregates even when they were composed of 90% of the control cells and 10% of the *Zeb2* KO ESCs (Fig. 30H,I). However, after confocal microscopy examination of the aggregates, we discovered that a small percentage of Zeb2KO-GFP ESCs differentiated into BIII-Tubulin+ neurons as indicated by a double GFP/BIII-Tubulin IF staining (Fig. 30B,C). Interestingly, the majority of Zeb2KO-GFP cells that formed small clusters within the Ctrl/KO aggregates stayed undifferentiated as shown by Oct4 IF staining (Fig. 30H,I) even after 40 days of the *in vitro* co-culture (*data not shown*). We conclude that *in vitro* Zeb2 controls transition from ESC to neuron primarily in a cell-autonomous fashion.

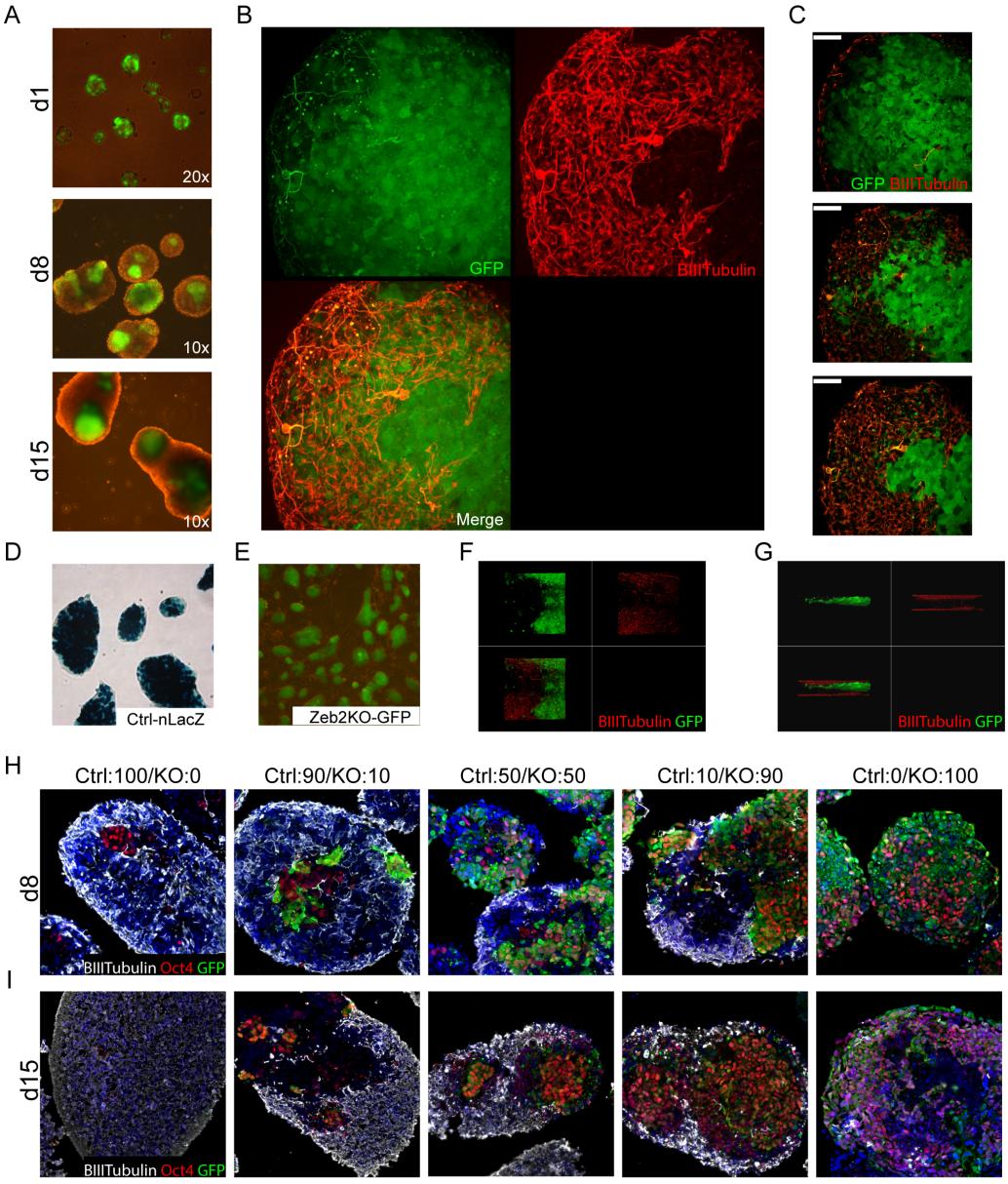

Figure 30. Zeb2 action *in vitro* is primarily cell-autonomous.

A. Ctrl:50/KO:50 embryoid bodies in culture on d1,d8,d15. B. Maximum projection confocal image of Ctrl:10/KO:90 embryoid body on d15 of ND. GFP (marking Zeb2KO-GFP cells) shown in green. BIII-Tubulin shown in red C. Confocal image of the same EB as in B.; section through three planes. D. X-gal staining of Ctrl-nLacZ ESCs in culture. E. Zeb2KO-GFP ESCs in culture. F. Z-stack confocal image of Ctrl:50/KO:50 embryoid body on d15 of ND. View from the top. G. Z-stack confocal image of Ctrl:50/KO:50 embryoid body on d15 of ND. View from the side. H. IF staining for GFP (green), Oct4 (red),BIII-Tubulin (white) and Dapi (blue) in the Ctrl/KO aggregates on d8 of ND. I. IF staining for GFP (green), Oct4 (red), BIII-Tubulin (white) and Dapi (blue) in the Ctrl/KO aggregates on d15 of ND.

4.3.11 High percentage Zeb2KO-GFP/CD1 chimeric embryos show severe defects in early embryonic development

The ability of ESCs to contribute to chimeric embryos and remain also germline transmission competent is one of the key tests for ESC pluripotency. This technique can also be used to study cell potency in generating various embryonic lineages. We hypothesized that *Zeb2* KO ESCs would contribute to the neural tissues of the chimeric embryo with lower efficiency than control ESCs, and the developing embryo would preferably use the wild-type CD1 ESCs to build *e.g.*. the central nervous system. To test the pluripotency status of the *Zeb2* KO ESCs and to verify if these cells could efficiently contribute to this, we performed chimera experiments using CD1 host/acceptor morulae and Zeb2KO-GFP ESCs. We used Ctrl-nLacZ ESCs as positive control cells.

We observed that within 24h (when the aggregation embryos reach the blastocyst stage) Zeb2KO-GFP ESCs were efficiently incorporated in the chimeric embryos. Interestingly, Zeb2-deficient ESCs preferably contributed to the ICM (Fig. 31A). However, dissection at E9.0 revealed that only low percentage Zeb2KO-GFP/CD1 chimeric embryos could develop properly (only one low-percentage chimera was obtained, see Fig.31B). High-percentage chimeric embryos were severely underdeveloped (we obtained 8 severely underdeveloped chimeric embryos; See: Fig. 31B). In contrast, aggregation with the control ESC line yielded high numbers of well-developed high-percentage chimeric embryos (11 high-percentage and one low-percentage normally-developed chimera; See Fig. 31C).

The high-percentage Zeb2KO-GFP/CD1 chimeric embryos show severe developmental defects similar to those of *Zeb2* KO embryos (Van de Putte et al., 2003). In the chimera experiments the defects are even more pronounced as many of the high-percentage chimeras fail to form a proper embryo (see Fig. 31B, left picture). On the other hand, in low-percentage chimeras (Fig. 31B, right picture), GFP+ cells can be found in various tissues of the developing mouse embryos indicating that Zeb2-deficient ESCs have the pluripotent potential. Nevertheless, Zeb2 is required for proper development.

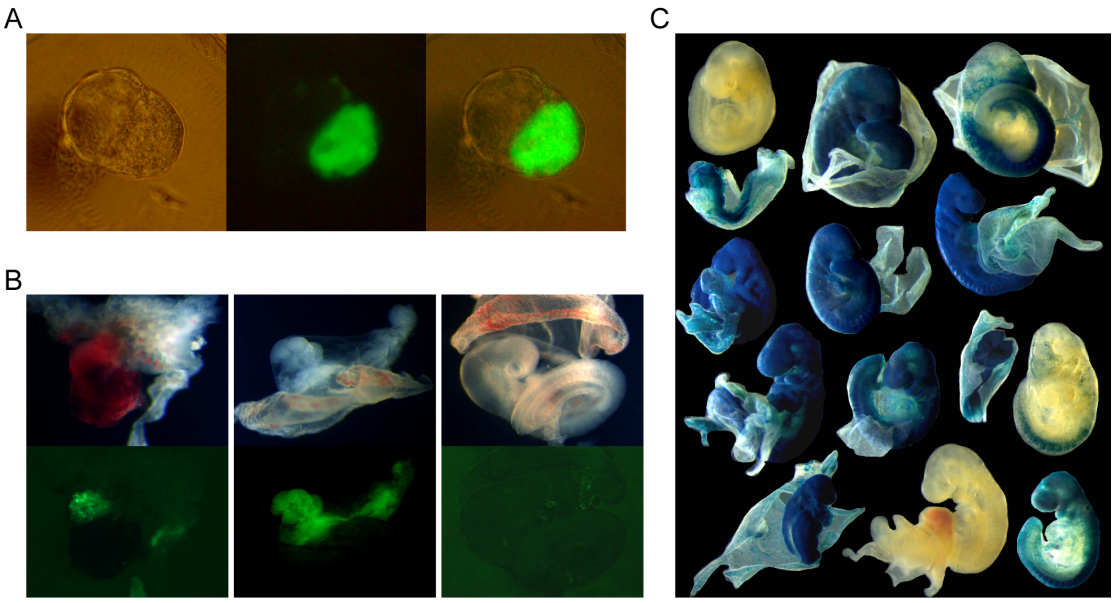

Figure 31. Zeb2 is essential for normal embryonic development.

A. Zeb2KO-GFP/CD1 blastocyst. B. Zeb2KO-GFP/CD1 chimeras. Only chimeras with low percentage of Zeb2KO-GFP cells are able to develop properly. C. Ctrl-nLacZ/CD1 chimeras develop normally (LacZ staining in blue).

4.3.12 Studies on R26-based Zeb2 domain mutant lines suggest that interactions of Zeb2 with some of its protein partners may have an inhibitory function on *in vitro* neural differentiation

Zeb2 is a multi-domain DNA-binding TF that interacts with a number of protein partners. The latter include activated Smads, the NuRD complex, CtBP co-repressor, and p300 and pCAF activators with HAT activity (Postigo *et al.*, 2003; Verschueren *et al.*, 1999; Verstappen *et al.*, 2008). In addition, our laboratory has done extensive proteomics, as well as post-translational modification studies in different cell systems and treatment conditions (with TGFβ or BMP), hence the list of candidate Zeb2 partners has grown significantly (Conidi *et al.*, *unpublished results*). Zeb2 and Zeb1 show 43% homology at protein level, however the homology is very high within the zinc fingers of both zinc finger clusters (88% for the NZF and 93% for the CZF cluster). In our hands, Zeb1 does not detectably bind to activated Smads like Zeb2 does via its linear Smad-binding domain (SBD), wherein 4 amino acid are crucial for Smad interaction (Conidi et al., 2013). Instead, based on work with fragments of Zeb1 it seems to require binding of p300 facilitating the formation of a Zeb1/p300/Smad activating complex (Postigo *et al.*, 2003).

Interestingly, early work in *Xenopus* demonstrated that both Zeb1 and Zeb2 interact with transcriptional co-activators p300 and pCAF and co-repressor CtBP-1, indicating that both Zeb TFs can

repress certain sets of genes, whilst activating other sets (van Grunsven *et al.*, 2006). To understand how Zeb2 functions in pluripotency and in (neural) development I started to generate in collaboration with the Gent-based members of the Haigh team (Gent, now Melbourne), *i.e.* Steven Goossens and Tim Pieters, a set of specific Zeb2 domain mutants expressed from the *R26* locus (under the control of endogenous *R26* promoter) in the *Zeb2* KO background (Haenebalcke *et al.*, 2013a). So far, the following mutants served these rescue attempts: mZeb2 wild-type (R26_Zeb2) (Fig. 32A), hZEB2 wild-type (R26_ZEB2) (Fig. 32C) and mZeb2 with a 153bp in-frame deletion of the Smad-binding domain (R26_Zeb2_dSBD) (Fig. 32D). In addition, we have also inserted mZeb1 in the *R26* locus of the *Zeb2* KO ESCs (Fig. 32B).

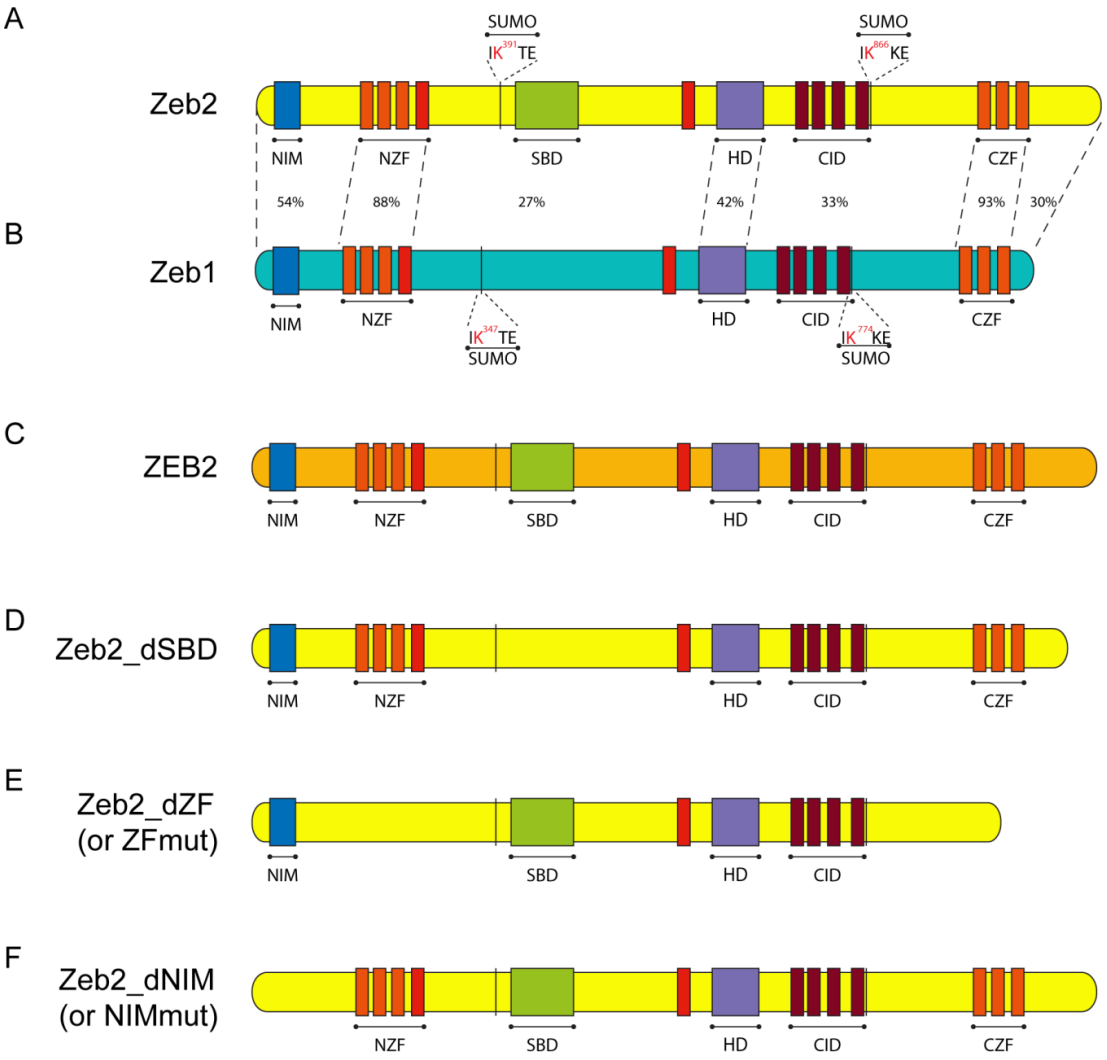

Figure 32. Schematic representation of mouse and human Zeb2 and mouse Zeb1 and their domain mutants.

NZF, CZF: Zinc finger clusters, CID: CtBP-interacting domain, SBD: Smad-binding domain, NIM: NuRD-interacting domain, HD: Homeodomain-like domain, SUMO: Sumoylation sites A. Wild-type mouse Zeb2 B. Wild-type mouse Zeb1 C. Wild-type human ZEB2 D. Mutant mouse Zeb2 with deletion of Smad-binding domain E. Mutant mouse Zeb2 with deletion of zinc fingers at N- and C-terminus F. Mutant mouse Zeb2 with deletion ofNuRD-interacting domain.

We subjected the *R26*-based rescue/KO combinatorial lines (Fig. 33A-D) to ND (3 clones per mutant line). Both R26_Zeb2 and R26_ZEB2 mutants partially restored the neural differentiation deficit caused by the *Zeb2* KO background (Fig. 33), and *R26*-based Zeb2 production partially silenced the persisting pluripotency program in those cells (Fig. 34). Neural differentiation rescue was more efficient in R26_Zeb2_dSBD and R26_Zeb1 in the *Zeb2* KO background indicating that Zeb2-Smad interaction might have an inhibitory function on Zeb2-driven neurogenesis, possibly by inhibiting inhibitory actions from residual activated Smads in neural development (Fig. 33).

We also verified Zeb2, Zeb1 and the neuronal marker Map2 presence on d0 and d15 in all *R26* mutant lines and we compared them to the control and *Zeb2* KO lines (Fig. 35 A,B,C). Zeb2 mRNA levels in the control pluripotent ESCs is very low and increase significantly (see elsewhere): the ESCs acquire neural fate and high expression on d15. *Zeb2* KO and R26_Zeb1 ESCs are negative for Zeb2. Zeb2 mRNA expression in R26_Zeb2, R26_ZEB2 and R26_Zeb2_dSBD was elevated already on d0 (approximately 50x higher than control cells on d0) since the *R26* locus is active in ESCs. Transgene expression (Zeb2 and Zeb1, respectively) from the *R26* locus remained unchanged in R26_Zeb2 and R26_ZEB2 lines and was elevated in the R26_Zeb2_dSBD and R26_Zeb1 lines. Theoretically, the *R26* locus is ubiquitously expressed in all tissues during development, but it is known to be a modesty active promoter in the brain and neuronal cells. Our RT-qPCR data shows that transgene expression from the *R26* locus was slightly higher in the lines that underwent neurogenesis most efficiently (R26_Zeb2_dSBD and R26_Zeb1, See Fig. 35A,B,C), suggesting that the activity of this locus is higher in neural cell types as compared to undifferentiated ESCs.

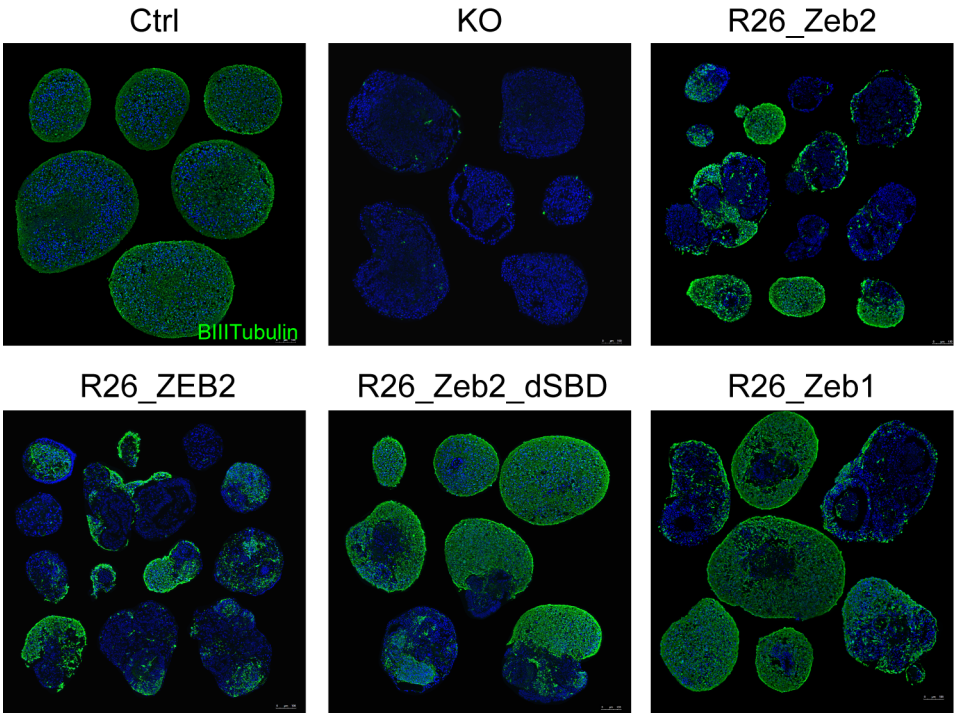

Figure 33.Neural differentiation of R26-based Zeb2 mutant lines.

Ctrl, Zeb2 KO and domain mutants were subjected to neural differentiation for 15 days and stained for BIII-Tubulin (green). Amongst R26 mutants, the best neural differentiation rescue was observed for R26_Zeb2_dSBD and R26_Zeb1 lines

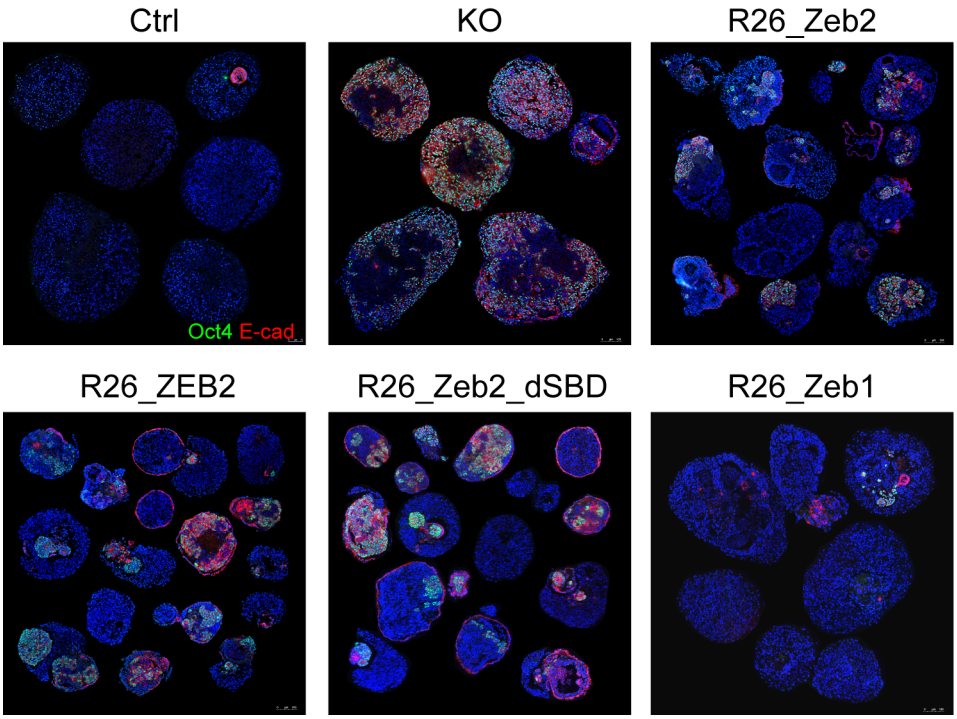

Figure 34. Neural differentiation of R26-based Zeb2 mutant lines.

Ctrl, Zeb2 KO and domain mutants were subjected to neural differentiation for 15 days and stained for Oct4 (green) and E-cadherin (red). Amongst the R26 mutants, the lowest E-cadherin and Oct4 levels were observed in R26_Zeb1 EBs.

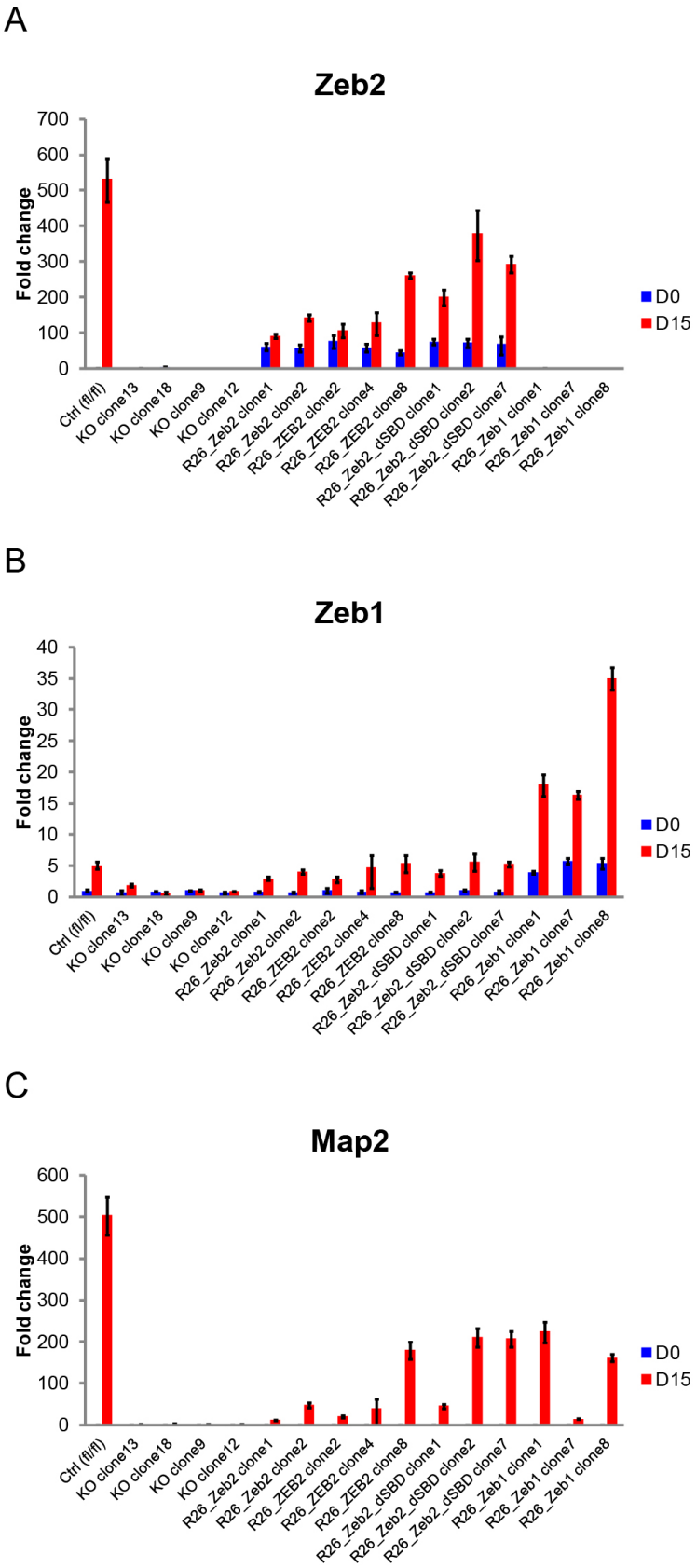

Figure 35. Neural differentiation of R26-based Zeb2 mutant lines.

RT-qPCR analysis of: A. Zeb2 B. Zeb1 C. Map2 mRNA expression on d0 (blue) and d15 (red) of neural differentiation.

4.4 Discussion

In Chapter 4 we addressed the role of Zeb2 in pluripotency maintenance as well as in pluripotency exit and ESC differentiation. Below, we separately discuss the main aspects of Zeb2 biology studied in the context of mouse ESCs.

4.4.1 Zeb2 is critical for exit from the epiblast state in ESCs and links the pluripotency network and DNA-methylation with irreversible commitment to differentiation.

A summary of the *Zeb2* KO phenotype and research highlights are shown in Fig. 36.

Using a knockout approach for *Zeb2* in ESCs for the first time with a rescue involving re-introduction of R26-locus driven Zeb2 cDNA, we show that Zeb2 is critical for these cells to undergo three-lineage differentiation. We propose that Zeb2 drives lineage commitment and specification by its action on multiple sets of Zeb2-dependent genes. First, Zeb2 is an important inducer of EMT (Comijn *et al.*, 2001; Vandewalle *et al.*, 2005). *Zeb2* KO ESCs retain epithelial characteristics when subjected to differentiating conditions. Their overall phenotype appears even more severe than the recently described knockout in ESCs of another known EMT regulator, *Snai1*, which still differentiate (Lin *et al.*, 2014). Second, the downregulation of important pluripotency network regulators depends on Zeb2. In contrast to Ctrl cells, *Zeb2* KO ESCs retain high Tet1, Oct4 and Nanog during differentiation. In ESCs, Tet1 is involved in a positive regulatory loop with Nanog and Oct4. Tet1 co-operates with Nanog, while a KD of Nanog weakens Tet1-binding to its target genes (including Oct4, Esrrb). Tet1 was also shown to act downstream of Oct4, and downregulation of Oct4 leads to decreased Tet1 expression (Ito *et al.*, 2010; Koh *et al.*, 2011; Costa *et al.*, 2013; Gao *et al.*, 2013). Tet1 was recently shown to act downstream of histone deacetylase Sirt6 to control ESC fate in differentiating conditions. Also, like in our system, the KD of Tet1 allowed silencing of *Oct4* and *Nanog* and rescued the *Sirt6* KO differentiation phenotype (Etchegaray *et al.*, 2015). Thus, Tet1 has a global inhibitory role in regulating *a.o.* key pluripotency genes during ESC differentiation and this work identifies Zeb2 as an (indirect) upstream factor important for achieving correct Tet1 levels.

We also describe a link between Zeb2 and regulation of DNA-methylation status. Acquisition of DNA-methyl marks during embryogenesis is thought to be unidirectional (Auclair *et al.*, 2014), but studies in ground-state, naïve ESCs and EpiSC, respectively, show that the methylomes are

interconvertible *in vitro* when different culture conditions are applied (Habibi *et al.*, 2013; Hackett *et al.*, 2013). Our RRBS showed that correct DNA-methylation patterns are initially acquired by *Zeb2* KO cells, but that this pattern cannot be sustained: indeed, *Zeb2* KO cells revert the methylome to a more naïve state, which agrees with the maintenance of their undifferentiated phenotype associated with persistence of Nanog and Oct4. Remarkably, this reversion in *Zeb2* KO cells is facilitated in absence of additional cues or signals, such as LIF or 2i. We hypothesize that Tet1 levels are maintained by the key pluripotency genes in *Zeb2* KO cells. Steady-state high-Tet1 would then actively demethylate the genome and contribute to preserving high Nanog and high Oct4 in the mutant cells. Enrichment of Tet1-binding at regions that lost methylation in the *Zeb2* KO further supports this hypothesis. Tet1 KD in these *Zeb2* KO cells facilitated downregulation of *Nanog* and *Oct4* as well as of *Cdh1*, but the differentiation phenotype was only partially rescued. It is likely that Zeb2 controls other important cell fate regulators at multiple stages of differentiation in addition to selected pluripotency genes and *Tet1*. Similar to the described *in vivo* functions of Zeb2 in myelinogenesis in the embryonic CNS (Weng *et al.*, 2012), Zeb2 may also counteract genes that are inhibitory for neural conversion during ESC differentiation and also act as an activator of other target genes depending on its co-factors (our unpublished data; Conidi *et al.*, 2011; Hegarty *et al.*, 2015), which altogether would then promote neurogenesis. As the current anti-Zeb2 antibodies do not allow omics-type studies (*not shown*), further more detailed studies will require an endogenous tagging approach to identify its protein partners and binding sites at multiple stages of differentiation.

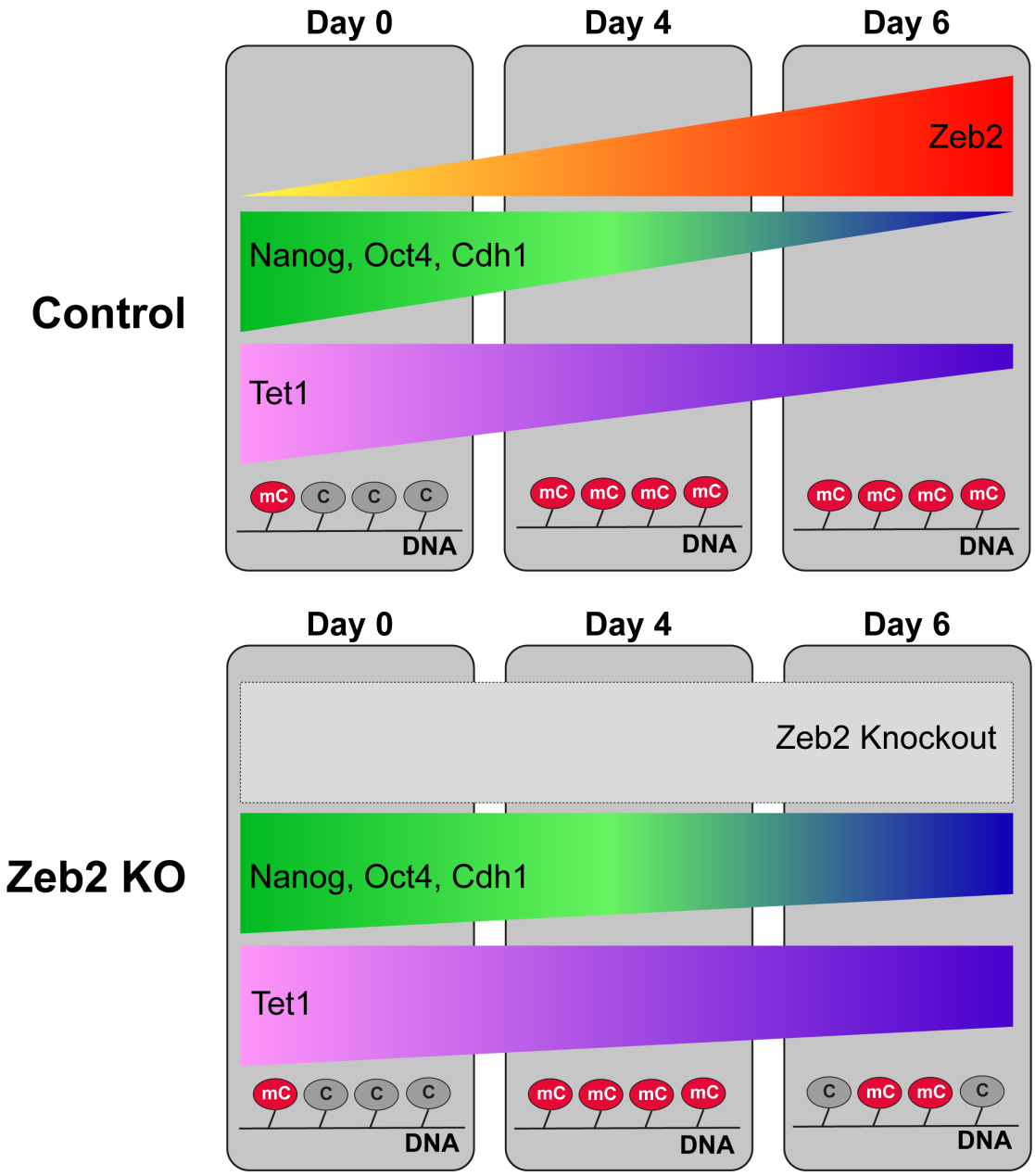

Figure 36. Graphical abstract: Zeb2 KO phenotype summary and highlights.

- *Zeb2* knockout (KO) ESCs display impaired differentiation in embryoid bodies by stalling in an epiblast-like state.
- In differentiating conditions EMT, pluripotency, lineage commitment and DNA-(de)methylation genes are deregulated in *Zeb2* KO embryoid bodies.
- *Zeb2* KO embryoid bodies cannot maintain their initially acquired DNA-methylation marks in neural-stimulating condition.
- *Zeb2* KO embryoid bodies do not effectively downregulate *Oct4*, *Nanog* and *Tet1* in differentiation conditions; Tet1 knockdown partially rescues their impaired differentiation.

4.4.2 Zeb2 and Rest

Rest is abundant in ESCs in pluripotency-supporting conditions and it is known for its anti-neurogenesis actions. It was also shown to be directly activated by Nanog, and depletion of Nanog results in a decrease of Rest levels (Loh *et al.*, 2006). Our temporal RNA-seq revealed that *Rest* was not properly downregulated in the *Zeb2* KO EBs. We reasoned that repression of an inhibitory gene for neural differentiation could be an additional mode of action by which Zeb2 controls neural lineage commitment and differentiation.

Knockdown of Rest in the Zeb2-deficient ESCs indeed restored some of their neurogenic capacity. Interestingly, forced downregulation of Rest did not cause a significant loss of pluripotency gene expression (unlike for Tet1 KD, which resulted in efficient downregulation of *Nano*g and *Oct4*). We hypothesize that knockdown of Rest in the *Zeb2* KO ESCs was able to rescue neural differentiation only in a small population of these cell cultures and in which the pluripotency genes were already destabilized and, hence, these cells were able to respond to the permissive conditions caused by knockdown of Rest.

Our data shows that Zeb2 regulates neurogenesis via at least two action modes: on the one hand it is critical for the correct control of Tet1 levels and on the other hand Zeb2 ensures efficient downregulation of *Rest*. It may be that *Nanog* (which we believe is a direct target of Zeb2) is the common upstream regulatory TF in these two modes and that the primary function of Zeb2 is to repress *Nanog*. It would be interesting to combine the effect of Tet1 and Rest KDs in *Zeb2* KO ESCs to see if *Zeb2* KO cells forced out of their pluripotent state by downregulation of Tet1, would still efficiently respond to Rest KD and therefore enter neural differentiation. In addition, generating Nanog KD ESC lines in the *Zeb2* KO background to investigate the phenotype rescue would also be an interesting route.

4.4.3 Cell-autonomous action of Zeb2 *in vitro*

Extrinsic factors are important regulators of cell fate specification (see General Introduction and Chapter 3 introduction). We asked whether Zeb2 could act in cell non-autonomous way during ESC differentiation. For this we performed co-culture differentiation experiments using differentially labeled *Zeb2* KO and Ctrl ESCs. We observed that Zeb2 action *in vitro* is primarily cell-autonomous as the *Zeb2* KO cells to a large extend maintained their undifferentiated phenotype even when surrounded by the

Ctrl cells. This is in contrast with our data on teratomas: there, *Zeb2* KO ESCs subcutaneously injected into the recipient mice underwent very efficient differentiation to all three lineages.

It may be that signaling molecules secreted by the Ctrl cells *in vitro* during differentiation are not produced in sufficient quantities and/or are too diluted in the culture medium. It may also be that extrinsic factors specifically needed for the cells to silence their pluripotency program, are abundantly present *in vivo* in adult tissues: *Zeb2* KO ESCs are then forced to respond to these signals, but these are absent (or nearly absent) from the co-culture system. If that is the case, even if the pro-neural factors (or in general pro-differentiation factors) are produced by the Ctrl cells, they can only act on neural progenitors but not on pluripotent cells, thus the whole, sequential fate decision program collapses.

4.4.4 Zeb2 and Smads

We analyzed the function of Zeb2 SBD during neural differentiation of ESCs. We first showed that in *Zeb2* KO cells in which Zeb2 wild-type cDNA is expressed from the *R26* locus the neural differentiation capacity is restored. We also demonstrated that insertion of ZEB2 wild-type cDNA results in the phenotype rescue, which shows both specificity as well as conservation of the Zeb2 function between mouse and human. Next, we demonstrated that Zeb2 mutant cDNA lacking the entire SBD (51 amino acids) inserted in the *R26* locus restored neural differentiation capacity in the *Zeb2* KO cells more efficiently than the wild-type Zeb2. This indicates that Zeb2-Smad co-operation, likely via direct interaction, during the transition from an ESC to a neural progenitor, has an inhibitory effect on neurogenesis.

This is in contrast with our previous data addressing the role of Zeb2-SBD in the ventral forebrain. We showed that the Zeb2-SBD is crucial for Zeb2-dependent interneuron migration in the embryonic brain. Embryonic brain slices, focally electroporated with a WT Zeb2 expression construct, restored partially interneuron migration, whereas this process could not be rescued when the SBD deletion construct was used (Conidi *et al.*, 2013).

Recently, Zeb2 was shown to govern myelinogenesis from oligodendrocyte precursor cells in te CNS by direct modulation of two Smad signaling pathways (Weng *et al.*, 2012). Activation of Bmp-Smads has an inhibitory effect on oligodendrocyte specification, while an increase of Smad7 (an inhibitory Smad) is required for oligodendrocyte myelinogenesis. Zeb2 interacts with phosphorylated Smad1 in the presence of the transcriptional co-activator p300 in the CNS. Zeb2 was also shown to bind directly to *Id2* and *Id4* promoters, which suggests direct transcriptional inhibition. Overproduction of

transduced Zeb2 in Bmp-stimulation conditions reverses Bmp-induced suppression of oligodendrocyte precursor genes known to drive myelinogenesis. Moreover, Smad7 promotes oligodendrocyte maturation and is directly activated by Zeb2 via Zeb2-binding sites in the *Smad7* promoter. Another interesting observation is that action of Zeb2 and its binding to Smads is stage-dependent. Therefore, different Zeb2 co-factors may direct recruitment of Zeb2 to target genes at specific time-points. In summary, Weng *et al.* (2012) showed that Zeb2 is in the center of the protein/transcriptional network of oligodendrocyte-dependent myelinogenesis and its actions converge on Bmp-Smad modulation in myelinogenesis in the CNS (Weng *et al.*, 2012).

Smad proteins are the intracellular mediators of the meanwhile named canonical Tgfβ/Bmp signaling pathway. Smads (Smad1/5) have been shown to act together with the core factors in the pluripotency network of ESCs (Chen *et al.*, 2008), but they are dispensable for ESC self-renewal (Gomes Fernandes et al., 2016; Morikawa et al., 2016). In human ESCs, *ZEB2* is repressed by Smad 2/3, Oct4 and Nanog (Chng *et al.*, 2010). Upon LIF withdrawal ZEB2 levels increase. It may be that in mouse ESCs a similar mechanism exists. *Zeb2* first needs to be derepressed by removal of the key pluripotency genes and Smads from its promoter. Higher Zeb2 could in turn specifically recognize and translocate to sites occupied by Smad proteins (or Zeb2 could be first bound by Smads and next, as a complex, it would translocate to the target sites), for example enhancer and promoter regions of the core pluripotency factors. If Zeb2-binding sites are present on the Smad-bound genes then it could directly regulate target gene transcription or, alternatively, Zeb2 recruited to Smads could subsequently recruit the NuRD complex to fine-tune gene expression regulation.

Based on published data and our observation in the ESC system we conclude that Zeb2-Smad protein-protein as well as protein-DNA interactions and their significance of directing cell fate decisions depend on the cellular context and most probably on the availability of other TFs and transcriptional co-regulators. It is possible that this interaction can be inhibitory for the initial steps of ESC specification, whereas later during development, for processes like interneuron migration it could be necessary.

It would be interesting to develop tools that allow manipulation of Zeb2-Smad binding and subsequent readout of these interactions in real-time. This way one would be able to pinpoint exactly what steps of what processes depend on intact Zeb2-Smad binding.

5 Chapter 5: General Discussion

Cell fate decisions in pluripotency, during development and in disease are orchestrated by a large number of diverse factors including extrinsic signaling proteins, transcription factors, non-coding RNAs, long-range genomic interactions and various epigenetic modifiers. Together they form an extremely complex and dynamic, but highly coordinated network with interconnected regulatory loops and multi-level modulations. Our laboratory studies how Zeb2 functions in distinct processes in normal development and more recently extended these studies to adult tissues and organs, including in injury conditions where repair is likely dependent on re-activation of developmental pathways. Remarkably, Zeb2 was shown to play a role in a wide variety of cell types and tissues: among others it regulates blood cell specification, melanocyte formation, brain development and T-cell and NK cells in the immune system, and CNS myelinogenesis and PNS (re)myelination by Schwann cells (see Chapter 1, general introduction). Below, we discuss the latest data from our laboratory, with focus on Zeb2-depedent control of neurogenesis also beyond birth.

The main part of my PhD research aimed at contributing to our emerging understanding whether and how Zeb2 functions in pluripotency and during differentiation of stem cells, using ESCs as a model system. We have shown that in ESCs Zeb2 is necessary to silence the pluripotency program at least in part via Tet1-mediated mechanisms. Upon Tet1 knockdown, Zeb2-deficient cells exit from their epiblast-like state, however their differentiation capacity is not fully restored. We hypothesize that Zeb2-dependent cell fate decisions during ESC differentiation rely on required interaction and/or loss of interactions with partner proteins or complexes thereof including *e.g.*. the NuRD complex (and/or individual components of this complex such as Mbd3 or HDAC1), while other new candidate regulators such as O-GlcNAc transferase (Conidi *et al*., *unpublished results*) or miRs, like the well-studied mir-200 family that operates in multiple cell types. Below we discuss also selected published data and explain how such interactions could contribute to pluripotency maintenance or acquisition of differentiated cellular phenotypes that form the core of this PhD research.

5.1 Does Zeb2 regulate other processes during brain development and in the adult?

We have shown that in the embryonic brain Zeb2 acts via at least three different mechanisms, depending on the location in the developing brain: it fine-tunes Wnt signaling to control hippocampus

formation (Miquelajauregui *et al.*, 2007), in brain cortex Zeb2 acts in a cell non-autonomous fashion by limiting the expression of *Nt3* and *Fgf9,* which normally ensure proper timing and extend of neurogenesis and gliogenesis, respectively (Seuntjens *et al.*, 2009), while in the ventral forebrain Zeb2 cell-autonomously regulates *a.o. Unc5b* transcription levels to assure proper cortical interneuron guidance and fate (van den Berghe *et al*., 2013; McKinsey *et al*., 2013)

In our laboratory we have started to expand studies of Zeb2's functions in early postnatal and in adult neurogenesis (with focus on the generation and function of the SEZ stem cell niche (also named the SVZ niche), the rostral migratory stream and olfactory bulb formation and its organization and cell diversity. The latest data suggests that Zeb2 is a major player in the establishment in the embryo and subsequently the maintenance and function of his neurogenic niche in the adult animal. Neuroblasts (also named A-cells) that later will become mature olfactory bulb (OB) interneurons are indeed born in the SVZ of the lateral ventricles from B-cells there, via a transit-amplifying cell population (the C-cells). The SVZ niche develops mainly from the ventral forebrain, from precursor cells in the embryonic LGE. Zeb2 is indeed present both in cells derived from the LGE during development as well as in the postnatal SVZ. Genetic inactivation of *Zeb2* from LGE cells (using the Gsh2-Cre deleter mouse; Van den Berghe *et al.*, 2013) results later in severe decrease in numbers of cells migrating from the SVZ to the OB and disorganization of the different OB cellular layers as well as in an overall reduced size of the OB at postnatal day 5 (at birth, the OB of these cell-type specific *Zeb2* KO mice appears to develop normally; Stappers *et al.*, *unpublished results*). Only a portion of *Zeb2* KO cells can reach the OB for their subsequent maturation there. Interestingly, the non-targeted cells that reach the OB also acquire an abnormal phenotype (characterized *a.o.* by increased *Pax6* expression), suggesting that Zeb2 acts in a cell non-autonomous way, likely on top of its cell-autonomous actions, in this process. Electroporation of the full-length Zeb2 (cDNA-driven) expression construct in the postnatal SVZ (d3) restores the neuronal cell influx to the OB. Conversely, the *Zeb2* KO phenotype can be induced by electroporation of Cre-expression construct in the $Zeb2^{fl/fl}$ SVZ (Stappers, Seuntjens *et al.*, *unpublished results*). Preliminary RNA-seq data from Ctrl and *Zeb2* KO cells (from the lateral wall of the SVZ at postnatal day 3) suggests that in absence of Zeb2 SVZ progenitors are misspecified and that some of these cells may acquire striatal neuron fate (this fate is established during embryonic development) that manifests in the early postnatal life. Taken together, this data shows that Zeb2 controls production and/or migration of olfactory neuronal cells from the SVZ neurogenic niche. It would be interesting to investigate whether Zeb2 is also involved in adult neurogenesis, especially in mobilization of neurogenic progenitors upon brain injury. This could also be assessed in the same injury studies by transplanting traceable ESCs with *Zeb2* mutations, including our *Zeb2* KO ESCs in the first place.

5.2 Does the NuRD complex co-operate with Zeb2 in the establishment of the correct chromatin context in embryonic cells, including in Zeb2-dependent differentiation?

The main focus of my PhD research is on the role of Zeb2 in cellular pluripotency and in early cell fate decisions, hence in particular during transition from pluripotent cell to committed progenitor cells of (likely) all three lineages. Below, based on recently published studies and our data, I will discuss how Zeb2, Dnmts, Tet1 and the NuRD complex (previously identified as an interacting partner of Zeb2: Verstappen *et al.*, 2008) could orchestrate/execute cell fate decisions and induce heterogeneity within populations of ESCs.

We show that *Zeb2* KO ESCs maintain their self-renewal capacity in cell culture, including in medium without 2i. Because of the observed phenotype in differentiation, we reasoned that impairment in pluripotency exit was the primary cause of defective differentiation in the absence of Zeb2. In order for the ESCs to differentiate, their pluripotency program needs to be silenced. *Oct4* and *Nanog* undergo transcriptional silencing through DNA-methylation that is maintained upon cell division in somatic cells. DNA-methylation is catalyzed by Dnmt proteins. Incomplete methylation caused by ablation of *Dnmt3a and/or Dnmt3b* leads to aberrant silencing of *Oct4* and *Nanog* during cell differentiation and in the embryo (Li *et al.*, 2007b) and is still being actively studied in 2i to serum conversion (and vice versa) in ESCs in the EC-FP7 project Blueprint, coordinated by H. Stunnenberg (Nijmegen). In our differentiation systems the Dnmt3a levels were lower, late during differentiation of the *Zeb2* KO ESCs. It would be interesting to see if very high levels of Dnmt3a (via its overexpression) in the *Zeb2* KO background would be sufficient to counteract the high Tet1 levels and silence the pluripotency program in ESCs, allowing differentiation to proceed. With other words, this would contribute to the ongoing studies in other groups that document how the precise, but also dynamic balance between DNA-methylation and DNA-demethylation co-determines these decisions or not.

The NuRD complex plays a crucial role in retaining ESC identity as well as in promoting transcriptional heterogeneity and lineage commitment (Kaji *et al.*, 2006; Reynolds *et al.*, 2012). Methyl-binding CpG proteins (Mbd) 1-4 are components of the NuRD complex (see also General Introduction). Mbd1, 2 and Mbd4 recognize and bind to methylated cytosine, whereas Mbd3 does not (Hendrich and Bird, 1998; Zhang *et al.*, 1999). Tet1 and Mbd3 binding profiles strongly overlap in pluripotent ESCs. Mbd3-bound target genes are enriched for the ^{5hm}C mark and knockdown of Tet1 leads to delocalization of Mbd3 binding to its genomic sites, while the total Mbd3 levels do not change. This suggests that Mbd3 recognizes and preferentially binds to hydroxylated regions of the genome and that this process is Tet1-dependent. Interestingly, Mbd3 knockdown resulted in lower ^{5hm}C levels, showing that Mbd3

also has a global function in establishing and/or maintaining global methylation patterns (Yildirim *et al.*, 2011). Our *Zeb2* KO cells retain high levels of Tet1. If high Tet1 and hence DNA-hydroxylation is abnormal in *Zeb2* KO ESCs, then Mbd3 recruitment to its genomic sites could also be defective. This could also contribute to aberrant gene expression regulation.

Mbd3 is a key component of NuRD (Kaji *et al.*, 2006). It would be of interest to evaluate genome-wide NuRD-dependent binding of Zeb2 to its target genes in pluripotency and during differentiation and to document the overlap with Tet1-binding and changes in ^{5hm}C levels. Here, an ESC line carrying the NuRD-interaction motif (NIM) of Zeb2 would be invaluable. As mentioned before, Zeb2 is very likely to operate in various modes in pluripotency exit and cell differentiation; its function(s) largely would then depend on its protein partners (as shown and discussed above for Zeb2-Smad binding). Zeb2 could - via repression of *Nanog* and other genes connected to the pluripotency network - regulate Tet1 levels and hence ^{5m}C/^{5hm}C ratios and the redistribution/dynamics of these marks in the genome. ^{5m}C/^{5hm}C redistribution and ratios may dictate which proteins and protein complexes can bind to specific loci. Zeb2 could partner with NuRD to directly regulate a number of target genes, it could also be guided by the NuRD complex (or Mbd3 alone, which remains to be tested) to the genomic sites marked by ^{5hm}C to selectively target gene expression in these loci. Thus we hypothesize that Zeb2-NuRD interaction may have a profound role specifically in ^{5hm}C-dependent gene expression regulation in cultured, including differentiation-induced ESCs, in particular in neural development.

In addition to Mbd3-Tet1 interaction, an alternative mechanism could contribute to the aberrant epigenetic status of Zeb2-deficient ESCs: HDACs remove acetyl marks from the histones resulting in a more closed chromatin. Hdac1 and Hdac2 are components of the NuRD complex. Knockdown of Hdac2 during reprogramming promotes iPSCs maturation. Costa *et al.* showed that knockdown of Hdac2 (but not Hdac 1,3 or 8) results in an increase in acetylation of histones. This increase facilitates Tet1-binding to *Nanog*, *Oct4*, *Sall4* and *Tcl1* and subsequent DNA hydroxylation (Wei *et al.*, 2015). It was previously shown that Tet1 binding depends on *Nanog* expression (Costa *et al.*, 2013). In line with the study by Costa and co-workers showed that the knockdown of Nanog in control or Hdac knockdown ESC lines resulted in a downregulation of Tet1 and a decrease in DNA hydroxylation. Knockdown of Tet1 in the Hdac knockdown cells blocked iPSC formation. RbAp46 (another component of NuRD) binds to both Tet1 and Hdac2. Upon RbAp46 knockdown, iPSC formation was abolished. The authors show that Hdac2 is bound to maturation phase-related gene promoters (*Nanog*, *Esrrb*, *Tcl1*, *Rex1*) in pre-iPSCs, whereas the same promoter region was occupied by Tet1 in fully reprogrammed iPSCs. Further, Hdac2 and Tet1 compete for binding to these regions (Wei *et al.*, 2015). In our *Zeb2* KO ESCs Tet1 is abundantly present even after LIF withdrawal. We hypothesize that Tet1 occupancy on the DNA is reinforced/stabilized by the persistently high Nanog levels. This strong binding does not allow Hdac2 to

access these target sites and would keep the chromatin open and hence prone for transcriptional activation.

5.3 Could high levels of Tet1 increase/stabilize recruitment of Ogt in the absence of Zeb2?

O-GlcNAcylation is addition of a single N-acetylglucosamine (GlcNAc) to Serine and Threonine residues in target proteins and (also) takes place in the nucleus (Haltiwanger *et al.*, 1992). Two enzymes catalyze the addition and removal, respectively, of *O*-GlcNAc in mammals: O-GlcNAc transferase (Ogt) adds the modification and O-GlcNAcase (Oga) removes it (for a review, see Hanover *et al.*, 2010). O-GlcNAcylation is known to influence protein function and stability. It is involved in nutrient (hence its levels are dependent on metabolic status, also because Ogt levels are *a.o.* controlled by 3 different metabolic pathways) and growth factor responses, cell cycle progression and cellular stress. One of the most important functions of O-glycosylation is to compete with and possibly prevent protein phosphorylation at specific sites (Zeidan and Hart, 2010).

Ogt is encoded by a single-gene throughout the animal kingdom, is essential for ESC viability and was shown to be a part of the pluripotency network in ESCs (Shafi *et al.*, 2000, van den Berg *et al.*, 2010; Pardo *et al.*, 2010). Short hairpin-mediated downregulation of Ogt levels results in decreased ESC proliferation and self-renewal. Conversely, increase in O-GlcNAcylation during EB formation inhibits their differentiation. Both Oct4 and Sox2 interact with Ogt and undergo O-GlcNAcylation that decreases rapidly once the ESCs are subjected to differentiation. O-GlcNAcylation regulates Oct4 transcriptional activity. Key genes of ESC self-renewal network such as *Klf2, Klf5, Nr5a2, Tbx3* and *Tcl1* are specifically regulated by O-GlcNAcylated Oct4. In addition, increasing O-GlcNAc levels by Ogt overexpression increased reprogramming efficiency (Jang *et al.*, 2012).

Tet1 directly binds to and is regulated by Ogt and can be O-GlcNAcylated. RNAi-mediated Ogt downregulation leads to a reduction in Tet1 and ^{5hm}C levels on Tet1 target genes (*e.g.*. *Lhx2* and *Pdgrfa*) and their de-repression. Conversely, ectopic expression of WT Ogt increases Tet1 levels (Shi *et al.*, 2013). Recently it was shown that Tet1 regulates Ogt recruitment to transcription start sites. Loss of Tet1 displaces Ogt binding from its target sites, which leads to an increase of common Tet1-Ogt target genes (Vella *et al.*, 2013). Interestingly, Ogt binding to DNA overlaps with Tet1 binding profiles at H3K4me3-marked, unmethylated CpG-rich gene promoters (Vella *et al.*, 2013). H3K4me3 was previously shown to “protect” CpGs from methylation by inhibiting Dnmts (reviewed in Hashimoto *et al.*, 2010). This suggests that Tet1-Ogt interaction plays a role in regulating promoter activity and gene

transcription regulation (rather than in moderation of DNA hydroxylation), which is consistent with the previously postulated dual function of Tet1 in (i) DNA hydroxylation and (ii) transcriptional regulation that is independent of Tet1 enzymatic activity (Dawlaty *et al.*, 2011; Williams *et al.*, 2011; Wu *et al.*, 2011; Xu *et al.*, 2011). It can be that Tet1 first recognizes and binds to unmethylated CpG regions and subsequently recruits various chromatin modifiers to regulate target gene expression. In the *Zeb2* KO ESCs Tet1 (stabilized by the core pluripotency network factors) would be strongly bound to the unmethylated cytosines together with Ogt. Abundance of Tet1 in absence of Zeb2 could alter cellular distribution of Ogt and its availability (amongst which a number of isoforms) for other proteins. In addition, Ogt could feedback to the pluripotency network and stabilize it. In fact, the situation is even more complex, as Zeb2 proteomics has identified Tet1 as a candidate partner protein and Zeb2 binds also Ogt and has lectin-affinity (Conidi *et al*., *unpublished results*).

Ogt could also be involved another type of epigenetic regulation. It was shown that Hdac1 is necessary for the full repressor activity of Sin3a (Hassig *et al.*, 1997). Both Hdac1 and Sin3a associate with Tet1 in ESCs (Shi *et al.*, 2013). Ogt co-purifies with Tet1 and Tet2, as well as Sin3a and Hdac1 (Vella *et al.*, 2013). Altogether this data suggests that Ogt is also a part of the Tet1/Tet2/Sin3a/Hdac1 complex in ESCs.

As mentioned above, our laboratory in a collaboration with E. Soler and F. Grosveld recently identified Ogt and Tet proteins as new candidate partners of Zeb2 (Conidi *et al*., *unpublished results*). Apart from its role in the pluripotency network, Ogt could also be important for assembly and/or guidance of various proteins or complexes that regulate the epigenetic status of the cells and hence their differentiation. It could be that Zeb2 - depending on the context - associates for example with NuRD, or acts together with Hdac1/Sin3a and Ogt (in smaller complexes then) to *e.g.*. silence gene expression. These interactions could depend both on the availability of the substrates as well as on the general chromatin structure that could potentially be altered by high levels of Tet1 in the *Zeb2* KO cells. It would be interesting to perform locus-specific proteomic screens (using targeted chromatin purification technique, like described by Pourfarzad *et al.*, 2013; or a CRISPR-based approach: Waldrip *et al.*, 2014) in ESCs as well as in differentiated cells to uncover when/how protein complexes are assembled and what the consequences are of their disruption in the context of Zeb2-dependent regulation of cell state and/or cell fate decisions.

5.4 Are Tet1 and miR-22 the new modulators of previously identified Zeb2-miR-200 feedback loop in ESCs?

Zeb2 is a major regulator of EMT (see General Introduction). EMT is one of the key events during ESC differentiation and in cancer progression (for a review, see Kim *et al.*, 2014). The miR-200 family was first shown to inhibit Zeb2 (and Zeb1) in cancer cell lines. In 60 cell lines tested, overexpression of miR-200 lead to upregulation of *E-cadherin* and subsequent reduced motility of the cells. Conversely, knockdown of miR-200 resulted in the induction of EMT and reduced *E-cadherin* levels (reviewed in: Korpal *et al.*, 2008; Park *et al.*, 2008).

Recently it was shown that miR-200 family members (miR-200a, miR-200b, miR-200c, miR-141, and miR-429) are mediators of OSKM-based reprogramming (Wang *et al.*, 2013). Exogenous Oct4 and Sox2 bind to the promoters of miR-141/200c and miR-200a/b/429 cluster, respectively, and induce expression of these miRs. Transcriptional activation of the miR-200 family induces MET by inhibiting *Zeb2*. In fact, forced downregulation of Zeb2 during iPSC generation results in a similar phenotype as miR-200 overexpression. Conversely, reprogramming is inhibited by overexpression of Zeb2 in the system. Together, this data shows that the mir-200--Zeb2 pathway is regulated by Oct4 and Sox2 and that modulation of this pathway is crucial for MET and acquisition of pluripotency (Wang *et al.*, 2013).

MiR-200 family expression is silenced by methylation of their promoters. In mammary gland tumors activation mir-200 is Tet1-dependent. Hydroxylation and activation of the miR-200 maintains the epithelial phenotype of the tumors hence reduces its aggressiveness and metastasis. Tet1 expression is inhibited by miR-22. Overexpression of miR-22 in human breast epithelial cell lines (HMEC and MCF-10A) or in non-metastatic breast cancer cell line (MCF-7) induced a mesenchymal phenotype and, after injection into the recipient mice, the MCF-7 line showed metastasis. MiR-22 directly represses Tet1 what leads to an increase in methylation of miR-200 promoters and their silencing. Thus, miR-22 is an important modifier of breast cancer stemness and metastasis and it exerts its function through Tet1-dependet regulation of miR-200 levels (Song *et al.*, 2013). miR-22 was previously identified as a miR whose expression increases during ESC differentiation (Houbaviy *et al.*, 2003).

Recently it was shown that Tet proteins are crucial for reprogramming to pluripotency. MEFs deficient in all three Tet proteins do not undergo MET and fail to generate iPSCs. This is due to lack of activation of the critical miR-200 family whose promoters stay methylated in absence of Tet proteins. Overexpression of Tet2 in the triple-*Tet* KO cells induced demethylation and partially restored miR-200 expression and reprogramming capacity. Interestingly, also overexpression of a single miR from the miR-200 family (i.e miR-200c) in the triple KO background allows iPSC generation. Together this data

show that Tet proteins are essential for reprogramming and they target the miR-200 family that is needed for MET transition during this process (Hu *et al.*, 2014).

We hypothesize that deregulation of the miR-200 regulatory loop in the *Zeb2* KO cells also reinforces their undifferentiated phenotype. Genetic inactivation of *Zeb2* removes the repressive marks from mirR-200. In addition, high levels of Tet1 retained in the *Zeb2* KO cells upon removal of 2i stabilize the active epigenetic state of miR-200 promoters. Moreover, Oct4, whose gene is not properly silenced in absence of Zeb2, promotes miR-200 expression. It would be interesting to investigate whether the Zeb2--miR-200-miR-22--Tet1 regulatory loop exists also in ESCs and whether manipulation of this pathway could have a beneficial effect on lineage-specific cell differentiation and reprograming.

Chapter 6: Global conclusions and perspectives

In my PhD research I studied two main aspects of Zeb2 biology.

In the first part, together with my colleagues, we specifically asked what the functions were of Zeb2 in certain aspects of mouse brain cortex development. For the first time in the field, we showed that Zeb2 controls expression of extrinsic factors in the upper layers of the cortex and that these deregulated factors feedback to the progenitors located in the (Zeb2-negative) ventricular zone to control cortex formation via controlling the timing of neurogenesis and gliogenesis.

In the second part we studied in cell culture how Zeb2 regulates cell fate decisions during progression from a pluripotent stem cell to a committed progenitor and beyond. We therefore documented mesodermal, endodermal and neuroectodermal lineage differentiation. Because of the mutations in *ZEB2* in humans lead to MOWS, we focused on the role of Zeb2 in transition from a ground state ESC to a neural progenitor. We discovered that Zeb2 regulates pluripotency exit at the epiblast stage in ESCs and that this process involves downstream action of Tet1. Analysis of the methylation status of Ctrl and *Zeb2* KO cells at three different time points during differentiation showed that *Zeb2* KO cells initially correctly acquire methyl marks on the DNA but they cannot maintain those marks and revert their methylome to a more naïve state, and that this loss is likely correlated with Tet1 binding. In addition we also initiated Zeb2 domain studies and we discovered that – unlike in cortical interneuron guided migration – the in-frame deletion of the Zeb2 SBD results in more efficient ESC differentiation towards neural cells. It would be interesting to expand our studies to other domains of Zeb2 and to investigate these interactions in various contexts of development and, subsequently, in disease.

Detailed knowledge of the dynamics of DNA binding of Zeb2, and its precise multi-modal role in regulation of gene transcription of specific sets of genes, is still lacking. Crucial will be to perform genome-wide and temporal ChIP-seq analysis of Zeb2-binding to its target loci. Based on published data on the role NuRD and the links between Tet1 and chromatin status (see General discussion), we hypothesize that Zeb2-deficient cells have also aberrant chromatin conformation. It would be interesting to document dynamic changes in the chromatin conformation during differentiation of Ctrl, *Zeb2* KO and Zeb2KO_Tet1shRNA ESCs. This new data, together with our RNA-seq and RRBS data, and the ChIP-

seq experiments would provide a more comprehensive picture of how Zeb2 functions in cell state and/or fate decisions.

It is known that averaging information from individual cells can often mask crucial information about fate decisions in biological systems. In the past years, technological advances took biomedical analytical experiments to the single-cell resolution. It is now possible to analyze the transcriptome from individual cells as well as investigate the epigenetic status of these (Faddah *et al.*, 2013; Kim *et al.*, 2015; Klein *et al.*, 2015; Kolodziejczyk *et al.*, 2015; Papatsenko *et al.*, 2015; Smallwood *et al.*, 2014). With such data, and notwithstanding the possible increasing confusion these floods of data may cause to most of us in the field, one may discover numerous distinct cell types, cell states and their origins in populations that thus far seem or are accepted to be relatively homogeneous (reviewed in Trapnell, 2015). Performing single-cell analysis in control and *Zeb2* KO cells in pluripotency and especially during early differentiation steps would provide additional valuable information on Zeb2's role in cell fate decisions. Our IHC/IF data shows that the majority of the cells in the *Zeb2* KO EBs after 15 days of differentiation are still highly positive for Oct4, Nanog and E-cadherin. RT-qPCR and RNA-seq expression (population) data indicates that (i) pluripotency gene expression is indeed not downregulated, (ii) there is an increase in early differentiation markers such as *Otx2* and *Fgf5*. Only single-cell resolution analysis could provide solid proof of the presence of two (or more) cell populations within the *Zeb2* KO EBs and whether these Nanog/Oct4/Ecad-negative cells are the actual EpiSCs or even more developmentally advanced types of cell. The same approach would also be very useful in answering the question about the cell non-autonomous function of Zeb2 in our co-culture experiment (see Chapter 4.3.9). Based on the GFP signal coming from the labeled Zeb2KO-GFP cells we concluded that only a very small percentage of ESCs lacking Zeb2 were able to acquire neural phenotype when co-cultured with Ctrl cells. However, from this data we could not exclude that some of the Zeb2KO-GFP cells silenced their GFP signal and we could not detect it using simple IF analysis. Single-cell analysis would enable us to profile the cells and correlate Zeb2 and GFP levels with neural marker gene sets and possibly demonstrate that also *in vitro* Zeb2 can induce neurogenesis in a cell non-autonomous fashion.

Integrating all levels of regulation (transcriptional, metabolic, epigenetic, extrinsic signals etc) in biological systems at a single-cell resolution will provide unbiased classification of cell types within organs, tissues and their *in vitro* counterparts as well as new definitions of previously unknown cellular states. It could also contribute to novel insights into plasticity of cell systems. Such data will provide information necessary to build to systems-biology models that would more accurately predict cell behavior and decisions during normal development, also upon their perturbation, and in disease.

Studies in mESCs and early embryos provided very valuable information about how pluripotency is established and maintained. It was shown that ESCs can be captured in ground state of self-renewal and that this state very strongly resembles E4.5 mouse blastocyst (Boroviak *et al.*, 2014; Ying *et al.*, 2008). Pluripotent mouse ESCs can also be propagated as EpiSCs in a more primed state that is equivalent to preimplantation epiblast (See General Introduction and: Brons *et al.*, 2007; Tesar *et al.*, 2007). Human ESCs under classic hESC culture conditions (Fgf2+Activin) resemble mouse epiblast stem cells. "Classic" hESCs show global DNA hypermethylation, relatively high expression of lineage-specific genes and Oct4 expression driven by its proximal enhancer. Recently a number of publications, inspired by the ground state of mouse ESC self-renewal, showed that naïve or possibly even ground state of pluripotency could be achieved in hESCs. Naïve state of human ESC self-renewal was obtained by transient gene overexpression (NANOG/KLF2 in presence of 2i+LIF: Takashima *et al.*, 2014; KLF2/KLF4 or OCT4/KLF4 in presence if 2i+LIF: Ware *et al.*, 2014) and/or by manipulation of extrinsic signaling pathways (using a comibination of:LIF, TGFβ1, FGF2, ERK1/2i, GSK3βi, JNKi, p38i: Gafni *et al.*, 2013; using 2i+LIF and Wnt5a: Van der Jeught *et al.*, 2015; using 2i+hLIF supplemented with Activin and ROCKi, BRAFi, SRCi: Theunissen *et al.*, 2014). Human ESCs after a relatively short transition period, acquire mESC-like features and in some cases show contribution to inter-species mouse-human chimeras (Gafni *et al.*, 2013).

Dissecting molecular mechanisms that underlay pluripotency and feed into the balance between maintenance or exit from self-renewal, defining and carefully documenting the spectrum of pluripotent states in mouse and human embryos as well as in ESCs and iPSCs is of highest importance as it broadens our knowledge about how developmental processes, cell differentiation and reprogramming are orchestrated and all this together is extremely relevant for research and biomedical purposes such as disease modeling, drug screening and cell therapy.

List of references

Abranches, E., Bekman, E., and Henrique, D. (2013). Generation and characterization of a novel mouse embryonic stem cell line with a dynamic reporter of Nanog expression. PLoS One *8*, e59928.

Aiba, K., Nedorezov, T., Piao, Y., Nishiyama, A., Matoba, R., Sharova, L. V, Sharov, A.A., Yamanaka, S., Niwa, H., and Ko, M.S.H. (2009). Defining Developmental Potency and Cell Lineage Trajectories by Expression Profiling of Differentiating Mouse Embryonic Stem Cells. DNA Res. *16* , 73–80.

Akalin, A., Kormaksson, M., Li, S., Garrett-Bakelman, F.E., Figueroa, M.E., Melnick, A., and Mason, C.E. (2012). methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. Genome Biol. *13*, R87.

Ang, Y.-S., Tsai, S.-Y., Lee, D.-F., Monk, J., Su, J., Ratnakumar, K., Ding, J., Ge, Y., Darr, H., Chang, B., et al. (2011). Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. Cell *145*, 183–197.

Arnold, S.J., and Robertson, E.J. (2009). Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. Nat Rev Mol Cell Biol *10*, 91–103.

Ashe, H.L., and Briscoe, J. (2006). The interpretation of morphogen gradients. Dev. *133* , 385–394.

Auclair, G., Guibert, S., Bender, A., and Weber, M. (2014). Ontogeny of CpG island methylation and specificity of DNMT3 methyltransferases during embryonic development in the mouse. Genome Biol. *15*, 545.

Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N., and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev. *17*, 126–140.

Bain, G., Ray, W.J., Yao, M., and Gottlieb, D.I. (1996). Retinoic acid promotes neural and represses mesodermal gene expression in mouse embryonic stem cells in culture. Biochem. Biophys. Res. Commun. *223*, 691–694.

Bao, S., Tang, F., Li, X., Hayashi, K., Gillich, A., Lao, K., and Surani, M.A. (2009). Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells . Nature *461*, 1292–1295.

Barnabé-Heider, F., and Miller, F.D. (2003). Endogenously Produced Neurotrophins Regulate Survival and Differentiation of Cortical Progenitors via Distinct Signaling Pathways. J. Neurosci. *23* , 5149–5160.

Bartolini, G., Ciceri, G., and Marín, O. (2013). Integration of GABAergic interneurons into cortical cell assemblies: lessons from embryos and adults. Neuron *79*, 849–864.

Bedzhov, I., Graham, S.J.L., Leung, C.Y., and Zernicka-Goetz, M. (2014). Developmental plasticity, cell fate specification and morphogenesis in the early mouse embryo. Philos. Trans. R. Soc. London B Biol. Sci. *369*.

van den Berg, D.L.C., Snoek, T., Mullin, N.P., Yates, A., Bezstarosti, K., Demmers, J., Chambers, I., and Poot, R.A. (2010). An Oct4-Centered Protein Interaction Network in Embryonic Stem Cells. Cell Stem Cell *6*, 369–381.

ten Berge, D., Kurek, D., Blauwkamp, T., Koole, W., Maas, A., Eroglu, E., Siu, R.K., and Nusse, R. (2011). Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells. Nat Cell Biol *13*, 1070–1075.

van den Berghe, V., Stappers, E., Vandesande, B., Dimidschstein, J., Kroes, R., Francis, A., Conidi, A., Lesage, F., Dries, R., Cazzola, S., et al. (2013). Directed migration of cortical interneurons depends on the cell-autonomous action of Sip1. Neuron *77*, 70–82.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315–326.

Betschinger, J., Nichols, J., Dietmann, S., Corrin, P.D., Paddison, P.J., and Smith, A. (2013). Exit from Pluripotency Is Gated by Intracellular Redistribution of the bHLH Transcription Factor Tfe3. Cell *153*, 335–347.

Boiani, M., and Scholer, H.R. (2005). Regulatory networks in embryo-derived pluripotent stem cells. Nat Rev Mol Cell Biol *6*, 872–881.

Borgel, J., Guibert, S., Li, Y., Chiba, H., Schubeler, D., Sasaki, H., Forne, T., and Weber, M. (2010). Targets and dynamics of promoter DNA methylation during early mouse development. Nat Genet *42*, 1093–1100.

Boroviak, T., Loos, R., Bertone, P., Smith, A., and Nichols, J. (2014). The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. Nat Cell Biol *16*, 513–525.

Boyer, L. a, Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947–956.

Brabletz, S., and Brabletz, T. (2010). The ZEB/miR-200 feedback loop—a motor of cellular plasticity in development and cancer? EMBO Rep. *11*, 670–677.

Bracken, C.P., Gregory, P.A., Kolesnikoff, N., Bert, A.G., Wang, J., Shannon, M.F., and Goodall, G.J. (2008). A Double-Negative Feedback Loop between ZEB1-SIP1 and the microRNA-200 Family Regulates Epithelial-Mesenchymal Transition. Cancer Res. *68* , 7846–7854.

Brons, I.G.M., Smithers, L.E., Trotter, M.W.B., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S.M., Howlett, S.K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R. a, et al. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature *448*, 191–195.

Buecker, C., Srinivasan, R., Wu, Z., Calo, E., Acampora, D., Faial, T., Simeone, A., Tan, M., Swigut, T., and Wysocka, J. (2014). Reorganization of enhancer patterns in transition from naïve to primed pluripotency. Cell Stem Cell *14*, 838–853.

Bultman, S., Gebuhr, T., Yee, D., La Mantia, C., Nicholson, J., Gilliam, A., Randazzo, F., Metzger, D., Chambon, P., Crabtree, G., et al. (2000). A Brg1 Null Mutation in the Mouse Reveals Functional Differences among Mammalian SWI/SNF Complexes. Mol. Cell *6*, 1287–1295.

Burgess, N., Maguire, E.A., and O'Keefe, J. (2002). The Human Hippocampus and Spatial and Episodic Memory. Neuron *35*, 625–641.

Burton, A., Muller, J., Tu, S., Padilla-Longoria, P., Guccione, E., and Torres-Padilla, M.-E. (2013). Single-Cell Profiling of Epigenetic Modifiers Identifies PRDM14 as an Inducer of Cell Fate in the Mammalian Embryo. Cell Rep. *5*, 687–701.

Cacheux, V., Dastot-Le Moal, F., Kääriäinen, H., Bondurand, N., Rintala, R., Boissier, B., Wilson, M., Mowat, D., and Goossens, M. (2001). Loss-of-function mutations in SIP1 Smad interacting protein 1 result in a syndromic Hirschsprung disease. Hum.

Mol. Genet. *10* , 1503–1510.
Cahoy, J.D., Emery, B., Kaushal, A., Foo, L.C., Zamanian, J.L., Christopherson, K.S., Xing, Y., Lubischer, J.L., Krieg, P.A., Krupenko, S.A., et al. (2008). A Transcriptome Database for Astrocytes, Neurons, and Oligodendrocytes: A New Resource for Understanding Brain Development and Function. J. Neurosci. *28* , 264–278.
Cai, C., and Grabel, L. (2007). Directing the differentiation of embryonic stem cells to neural stem cells. Dev. Dyn. *236*, 3255–3266.
Campos, E.I., and Reinberg, D. (2009). Histones: Annotating Chromatin. Annu. Rev. Genet. *43*, 559–599.
Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing. Sci. *298* , 1039–1043.
Cartwright, P., McLean, C., Sheppard, A., Rivett, D., Jones, K., and Dalton, S. (2005). LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism. Dev. *132* , 885–896.
Chambers, I., and Smith, A. (2004). Self-renewal of teratocarcinoma and embryonic stem cells. Oncogene *23*, 7150–7160.
Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell *113*, 643–655.
Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. Nature *450*, 1230–1234.
Chambery, A., Vissers, J.P.C., Langridge, J.I., Lonardo, E., Minchiotti, G., Ruvo, M., and Parente, A. (2009). Qualitative and Quantitative Proteomic Profiling of Cripto−/− Embryonic Stem Cells by Means of Accurate Mass LC−MS Analysis. J. Proteome Res. *8*, 1047–1058.
Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J., et al. (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell *133*, 1106–1117.
Chew, J.-L., Loh, Y.-H., Zhang, W., Chen, X., Tam, W., Yeap, L., Li, P., Ang, Y., Lim, B., Robson, P., et al. (2005). Reciprocal Transcriptional Regulation of Pou5f1 and Sox2 via the Oct4/Sox2 Complex in Embryonic Stem Cells. Mol. Cell. Biol. *25* , 6031–6046.
Chng, Z., Teo, A., Pedersen, R.A., and Vallier, L. (2010). SIP1 mediates cell-fate decisions between neuroectoderm and mesendoderm in human pluripotent stem cells. Cell Stem Cell *6*, 59–70.
Cinaroglu, A., Ozmen, Y., Ozdemir, A., Ozcan, F., Ergorul, C., Cayirlioglu, P., Hicks, D., and Bugra, K. (2005). Expression and possible function of fibroblast growth factor 9 (FGF9) and its cognate receptors FGFR2 and FGFR3 in postnatal and adult retina. J. Neurosci. Res. *79*, 329–339.
Comijn, J., Berx, G., Vermassen, P., Verschueren, K., van Grunsven, L., Bruyneel, E., Mareel, M., Huylebroeck, D., and van Roy, F. (2001). The Two-Handed E Box Binding Zinc Finger Protein SIP1 Downregulates E-Cadherin and Induces Invasion. Mol. Cell *7*, 1267–1278.
Conidi, A., Cazzola, S., Beets, K., Coddens, K., Collart, C., Cornelis, F., Cox, L., Joke, D., Dobreva, M.P., Dries, R., et al. (2011). Few Smad proteins and many Smad-interacting proteins yield multiple functions and action modes in TGF??/BMP signaling in vivo. Cytokine Growth Factor Rev. *22*, 287–300.
Conidi, A., van den Berghe, V., Leslie, K., Stryjewska, A., Xue, H., Chen, Y.-G., Seuntjens, E., and Huylebroeck, D. (2013). Four Amino Acids within a Tandem QxVx Repeat in a Predicted Extended α-Helix of the Smad-Binding Domain of Sip1 Are Necessary for Binding to Activated Smad Proteins. PLoS One *8*, e76733.
Cordelli, D.M., Garavelli, L., Savasta, S., Guerra, A., Pellicciari, A., Giordano, L., Bonetti, S., Cecconi, I., Wischmeijer, A., Seri, M., et al. (2013). Epilepsy in Mowat–Wilson syndrome: Delineation of the electroclinical phenotype. Am. J. Med. Genet. Part A *161*, 273–284.
Costa, Y., Ding, J., Theunissen, T.W., Faiola, F., Hore, T. a, Shliaha, P. V, Fidalgo, M., Saunders, A., Lawrence, M., Dietmann, S., et al. (2013). NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. Nature *495*, 370–374.
Custo Greig, L.F., Woodworth, M.B., Galazo, M.J., Padmanabhan, H., and Macklis, J.D. (2013). Molecular logic of neocortical projection neuron specification, development and diversity. Nat. Rev. Neurosci. *14*, 10.1038/nrn3586.
Dastot-Le Moal, F., Wilson, M., Mowat, D., Collot, N., Niel, F., and Goossens, M. (2007). ZFHX1B mutations in patients with Mowat-Wilson syndrome. Hum. Mutat. *28*, 313–321.
Dawlaty, M.M., Ganz, K., Powell, B.E., Hu, Y.-C., Markoulaki, S., Cheng, A.W., Gao, Q., Kim, J., Choi, S.-W., Page, D.C., et al. (2011). Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. Cell Stem Cell *9*, 166–175.
Dawlaty, M.M., Breiling, A., Le, T., Barrasa, M.I., Raddatz, G., Gao, Q., Powell, B.E., Cheng, A.W., Faull, K.F., Lyko, F., et al. (2014). Loss of Tet enzymes compromises proper differentiation of embryonic stem cells. Dev. Cell *29*, 102–111.
Debruyn, J. (2010). PhD Thesis.
Delcuve, G.P., Khan, D.H., and Davie, J.R. (2012). Roles of histone deacetylases in epigenetic regulation: emerging paradigms from studies with inhibitors. Clin. Epigenetics *4*, 5.
Denecker, G., Vandamme, N., Akay, O., Koludrovic, D., Taminau, J., Lemeire, K., Gheldof, A., De Craene, B., Van Gele, M., Brochez, L., et al. (2014). Identification of a ZEB2-MITF-ZEB1 transcriptional network that controls melanogenesis and melanoma progression. Cell Death Differ *21*, 1250–1261.
Desai, A.R., and McConnell, S.K. (2000). Progressive restriction in fate potential by neural progenitors during cerebral cortical development. Dev. *127* , 2863–2872.
Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380.
Dunn, S.-J., Martello, G., Yordanov, B., Emmott, S., and Smith, A.G. (2014). Defining an essential transcription factor program for naïve pluripotency. Sci. *344* , 1156–1160.
Eiraku, M., Watanabe, K., Matsuo-Takasaki, M., Kawada, M., Yonemura, S., Matsumura, M., Wataya, T., Nishiyama, A., Muguruma, K., and Sasai, Y. (2008). Self-organized formation of polarized cortical tissues from ESCs and its active manipulation by extrinsic signals. Cell Stem Cell *3*, 519–532.

Elloul, S., Bukholt Elstrand, M., Nesland, J.M., Tropé, C.G., Kvalheim, G., Goldberg, I., Reich, R., and Davidson, B. (2005). Snail, Slug, and Smad-interacting protein 1 as novel parameters of disease aggressiveness in metastatic ovarian and breast carcinoma. Cancer *103*, 1631–1643.

Emery, B. (2010). Regulation of Oligodendrocyte Differentiation and Myelination. Sci. *330* , 779–782.

Engelen, E., Akinci, U., Bryne, J.C., Hou, J., Gontan, C., Moen, M., Szumska, D., Kockx, C., van IJcken, W., Dekkers, D.H.W., et al. (2011). Sox2 cooperates with Chd7 to regulate genes that are mutated in human syndromes. Nat Genet *43*, 607–611.

Ernfors, P. (2010). Cellular origin and developmental mechanisms during the formation of skin melanocytes. Exp. Cell Res. *316*, 1397–1407.

Eskeland, R., Leeb, M., Grimes, G.R., Kress, C., Boyle, S., Sproul, D., Gilbert, N., Fan, Y., Skoultchi, A.I., Wutz, A., et al. (2010). Ring1B Compacts Chromatin Structure and Represses Gene Expression Independent of Histone Ubiquitination. Mol. Cell *38*, 452–464.

Etchegaray, J.-P., Chavez, L., Huang, Y., Ross, K.N., Choi, J., Martinez-Pastor, B., Walsh, R.M., Sommer, C.A., Lienhard, M., Gladden, A., et al. (2015). The histone deacetylase SIRT6 controls embryonic stem cell fate via TET-mediated production of 5-hydroxymethylcytosine. Nat Cell Biol *17*, 545–557.

Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154–156.

Faddah, D.A., Wang, H., Cheng, A.W., Katz, Y., Buganim, Y., and Jaenisch, R. (2013). Single-cell analysis reveals that expression of nanog is biallelic and equally variable as that of other pluripotency factors in mouse escs. Cell Stem Cell *13*, 23–29.

Fang, F., Xu, Y., Chew, K.-K., Chen, X., Ng, H.-H., and Matsudaira, P. (2014). Coactivators p300 and CBP Maintain the Identity of Mouse Embryonic Stem Cells by Mediating Long-Range Chromatin Structure. Stem Cells *32*, 1805–1816.

Faust, C., Schumacher, A., Holdener, B., and Magnuson, T. (1995). The eed mutation disrupts anterior mesoderm production in mice. Dev. *121* , 273–285.

Festuccia, N., Osorno, R., Wilson, V., and Chambers, I. (2013). The role of pluripotency gene regulatory network components in mediating transitions between pluripotent cell states. Curr. Opin. Genet. Dev. *23*, 504–511.

Festuccia, N., Osorno, R., Halbritter, F., Karwacki-Neisius, V., Navarro, P., Colby, D., Wong, F., Yates, A., Tomlinson, S.R.R., and Chambers, I. (2015). Esrrb Is a Direct Nanog Target Gene that Can Substitute for Nanog Function in Pluripotent Cells. Cell Stem Cell *11*, 477–490.

Fietz, S.A., and Huttner, W.B. (2011). Cortical progenitor expansion, self-renewal and neurogenesis—a polarized perspective. Curr. Opin. Neurobiol. *21*, 23–35.

Finlay, B.L., and Darlington, R.B. (1995). Linked regularities in the development and evolution of mammalian brains. Sci. *268* , 1578–1584.

Finley, M.F., Devata, S., and Huettner, J.E. (1999). BMP-4 inhibits neural differentiation of murine embryonic stem cells. J. Neurobiol. *40*, 271–287.

Flames, N., and Marín, O. (2005). Developmental mechanisms underlying the generation of cortical interneuron diversity. Neuron *46*, 377–381.

Franco, S.J., and Müller, U. (2013). Shaping our Minds: Stem and Progenitor Cell Diversity in the Mammalian Neocortex. Neuron *77*, 19–34.

Frantz, G.D., and McConnell, S.K. (1996). Restriction of Late Cerebral Cortical Progenitors to an Upper-Layer Fate. Neuron *17*, 55–61.

Fuda, N.J., Ardehali, M.B., and Lis, J.T. (2009). Defining mechanisms that regulate RNA polymerase II transcription in vivo. Nature *461*, 186–192.

Fukumitsu, H., Ohtsuka, M., Murai, R., Nakamura, H., Itoh, K., and Furukawa, S. (2006). Brain-Derived Neurotrophic Factor Participates in Determination of Neuronal Laminar Fate in the Developing Mouse Cerebral Cortex. J. Neurosci. *26* , 13218–13230.

Funahashi, J., Kamachi, Y., Goto, K., and Kondoh, H. (1991). Identification of nuclear factor delta EF1 and its binding site essential for lens-specific activity of the delta 1-crystallin enhancer. Nucleic Acids Res. *19*, 3543–3547.

Furutachi, S., Miya, H., Watanabe, T., Kawai, H., Yamasaki, N., Harada, Y., Imayoshi, I., Nelson, M., Nakayama, K.I., Hirabayashi, Y., et al. (2015). Slowly dividing neural progenitors are an embryonic origin of adult neural stem cells. Nat Neurosci *18*, 657–665.

Gafni, O., Weinberger, L., Mansour, A.A., Manor, Y.S., Chomsky, E., Ben-Yosef, D., Kalma, Y., Viukov, S., Maza, I., Zviran, A., et al. (2013). Derivation of novel human ground state naive pluripotent stem cells. Nature *504*, 282–286.

Gao, Y., Chen, J.J., Li, K., Wu, T., Huang, B., Liu, W., Kou, X., Zhang, Y., Huang, H., Jiang, Y., et al. (2014). Replacement of Oct4 by Tet1 during iPSC Induction Reveals an Important Role of DNA Methylation and Hydroxymethylation in Reprogramming. Cell Stem Cell *12*, 453–469.

Garavelli, L., and Mainardi, P.C. (2007). Mowat-Wilson syndrome. Orphanet J. Rare Dis. *2*, 42.

Gaspar-Maia, A., Alajem, A., Polesso, F., Sridharan, R., Mason, M.J., Heidersbach, A., Ramalho-Santos, J., McManus, M.T., Plath, K., Meshorer, E., et al. (2009). Chd1 regulates open chromatin and pluripotency of embryonic stem cells. Nature *460*, 863–868.

Gaspar-Maia, A., Alajem, A., Meshorer, E., and Ramalho-Santos, M. (2011). Open chromatin in pluripotency and reprogramming. Nat Rev Mol Cell Biol *12*, 36–47.

Gaspard, N., Bouschet, T., Herpoel, A., Naeije, G., van den Ameele, J., and Vanderhaeghen, P. (2009). Generation of cortical neurons from mouse embryonic stem cells. Nat. Protoc. *4*, 1454–1463.

Gelman DM, Marín; O, R.J. (2012). The Generation of Cortical Interneurons. In National Center for Biotechnology Information (US), pp. 1–13.

Germain, N., Banda, E., and Grabel, L. (2010). Embryonic stem cell neurogenesis and neural specification. J. Cell. Biochem. *111*, 535–542.

Gillich, A., Bao, S., Grabole, N., Hayashi, K., Trotter, M.W.B., Pasque, V., Magnúsdóttir, E., and Surani, M.A. (2012). Epiblast stem cell-based system reveals reprogramming synergy of germline factors. Cell Stem Cell *10*, 425–439.

Gillich, A., Bao, S., and Surani, M.A. (2013). Reversion of Mouse Postimplantation Epiblast Stem Cells to a Naïve Pluripotent State by Modulation of Signalling Pathways. In Epiblast Stem Cells SE - 2, R. Alberio, ed. (Humana Press), pp. 15–29.

Goll, M.G., and Bestor, T.H. (2005). EUKARYOTIC CYTOSINE METHYLTRANSFERASES. Annu. Rev. Biochem. *74*, 481–514.

Gomes Fernandes, M., Dries, R., Roost, M.S., Semrau, S., de Melo Bernardo, A., Davis, R.P., Ramakrishnan, R., Szuhai, K., Maas, E., Umans, L., et al. (2016). BMP-SMAD Signaling Regulates Lineage Priming, but Is Dispensable for Self-Renewal in Mouse Embryonic Stem Cells. Stem Cell Reports *6*, 85–94.

Goossens, S., Janzen, V., Bartunkova, S., Yokomizo, T., Drogat, B., Crisan, M., Haigh, K., Seuntjens, E., Umans, L., Riedt, T., et al. (2011). The EMT regulator Zeb2/Sip1 is essential for murine embryonic hematopoietic stem/progenitor cell differentiation and mobilization. Blood *117*, 5620–5630.

Goossens, S., Radaelli, E., Blanchet, O., Durinck, K., Van der Meulen, J., Peirs, S., Taghon, T., Tremblay, C.S., Costa, M., Ghahremani, M.F., et al. (2015). ZEB2 drives immature T-cell lymphoblastic leukaemia development via enhanced tumour-initiating potential and IL-7 receptor signalling. Nat Commun *6*.

Groszer, M., Erickson, R., Scripture-Adams, D.D., Lesche, R., Trumpp, A., Zack, J.A., Kornblum, H.I., Liu, X., and Wu, H. (2001). Negative Regulation of Neural Stem/Progenitor Cell Proliferation by the Pten Tumor Suppressor Gene in Vivo. Sci. *294* , 2186–2189.

van Grunsven, L.A., Michiels, C., Van de Putte, T., Nelles, L., Wuytens, G., Verschueren, K., and Huylebroeck, D. (2003). Interaction between Smad-interacting Protein-1 and the Corepressor C-terminal Binding Protein Is Dispensable for Transcriptional Repression of E-cadherin. J. Biol. Chem. *278* , 26135–26145.

van Grunsven, L.A., Taelman, V., Michiels, C., Opdecamp, K., Huylebroeck, D., and Bellefroid, E.J. (2006). δEF1 and SIP1 are differentially expressed and have overlapping activities during Xenopus embryogenesis. Dev. Dyn. *235*, 1491–1500.

Gu, T.-P., Guo, F., Yang, H., Wu, H.-P., Xu, G.-F., Liu, W., Xie, Z.-G., Shi, L., He, X., Jin, S., et al. (2011). The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. Nature *477*, 606–610.

Guo, J., and Anton, E.S. (2014). Decision making during interneuron migration in the developing cerebral cortex. Trends Cell Biol. *24*, 342–351.

Guo, G., Yang, J., Nichols, J., Hall, J.S., Eyres, I., Mansfield, W., and Smith, A. (2009). Klf4 reverts developmentally programmed restriction of ground state pluripotency. Dev. *136* , 1063–1069.

Gupta, A., Tsai, L.-H., and Wynshaw-Boris, A. (2002). Life is a journey: a genetic look at neocortical development. Nat Rev Genet *3*, 342–355.

Guttman, M., Amit, I., Garber, M., French, C., Lin, M.F., Feldser, D., Huarte, M., Zuk, O., Carey, B.W., Cassady, J.P., et al. (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature *458*, 223–227.

Guttman, M., Garber, M., Levin, J.Z., Donaghey, J., Robinson, J., Adiconis, X., Fan, L., Koziol, M.J., Gnirke, A., Nusbaum, C., et al. (2010). Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. Nat Biotech *28*, 503–510.

Guttman, M., Donaghey, J., Carey, B.W., Garber, M., Grenier, J.K., Munson, G., Young, G., Lucas, A.B., Ach, R., Bruhn, L., et al. (2011). lincRNAs act in the circuitry controlling pluripotency and differentiation. Nature *477*, 295–300.

Guzman-Ayala, M., Sachs, M., Koh, F.M., Onodera, C., Bulut-Karslioglu, A., Lin, C.-J., Wong, P., Nitta, R., Song, J.S., and Ramalho-Santos, M. (2015). Chd1 is essential for the high transcriptional output and rapid growth of the mouse epiblast. Development *142*, 118–127.

Habibi, E., Brinkman, A.B., Arand, J., Kroeze, L.I., Kerstens, H.H.D., Matarese, F., Lepikhov, K., Gut, M., Brun-Heath, I., Hubner, N.C., et al. (2013). Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. Cell Stem Cell *13*, 360–369.

Hackett, J.A., and Surani, M.A. (2014). Regulatory principles of pluripotency: from the ground state up. Cell Stem Cell *15*, 416–430.

Hackett, J.A., Dietmann, S., Murakami, K., Down, T.A., Leitch, H.G., and Surani, M.A. (2013). Synergistic mechanisms of DNA demethylation during transition to ground-state pluripotency. Stem Cell Reports *1*, 518–531.

Haenebalcke, L., Goossens, S., Naessens, M., Kruse, N., Ghahremani, M., Bartunkova, S., Haigh, K., Pieters, T., Dierickx, P., Drogat, B., et al. (2013a). Efficient ROSA26-Based Conditional and/or Inducible Transgenesis Using RMCE-Compatible F1 Hybrid Mouse Embryonic Stem Cells. Stem Cell Rev. Reports *9*, 774–785.

Haenebalcke, L., Goossens, S., Dierickx, P., Bartunkova, S., D'Hont, J., Haigh, K., Hochepied, T., Wirth, D., Nagy, A., and Haigh, J.J. (2013b). The ROSA26-iPSC mouse: a conditional, inducible, and exchangeable resource for studying cellular (De)differentiation. Cell Rep. *3*, 335–341.

Hall, J., Guo, G., Wray, J., Eyres, I., Nichols, J., Grotewold, L., Morfopoulou, S., Humphreys, P., Mansfield, W., Walker, R., et al. (2009). Oct4 and LIF/Stat3 additively induce Krüppel factors to sustain embryonic stem cell self-renewal. Cell Stem Cell *5*, 597–609.

Haltiwanger, R.S., Blomberg, M.A., and Hart, G.W. (1992). Glycosylation of nuclear and cytoplasmic proteins. Purification and characterization of a uridine diphospho-N-acetylglucosamine: polypeptide beta-N-acetylglucosaminyltransferase. J. Biol. Chem. *267*, 9005–9013.

Hanover, J.A., Krause, M.W., and Love, D.C. (2010). The hexosamine signaling pathway: O-GlcNAc cycling in feast or famine. Biochim. Biophys. Acta *1800*, 80–95.

Hashimoto, H., Vertino, P.M., and Cheng, X. (2010). Molecular coupling of DNA methylation and histone methylation. Epigenomics *2*, 657–669.

Hassig, C.A., Fleischer, T.C., Billin, A.N., Schreiber, S.L., and Ayer, D.E. (1997). Histone Deacetylase Activity Is Required for Full Transcriptional Repression by mSin3A. Cell *89*, 341–347.

Hayashi, K., de Sousa Lopes, S.M.C., Tang, F., Lao, K., and Surani, M.A. (2008). Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. Cell Stem Cell *3*, 10.1016/j.stem.2008.07.027.

Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S., and Saitou, M. (2011). Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. Cell *146*, 519–532.

Hayashi, K., Ogushi, S., Kurimoto, K., Shimamoto, S., Ohta, H., and Saitou, M. (2012). Offspring from Oocytes Derived from in Vitro Primordial Germ Cell–like Cells in Mice. Sci. *338* , 971–975.

Hegarty, S. V, Sullivan, A.M., and O'Keeffe, G.W. (2015). Zeb2: A multifunctional regulator of nervous system development. Prog. Neurobiol. *132*, 81–95.

Hendrich, B., and Bird, A. (1998). Identification and Characterization of a Family of Mammalian Methyl-CpG Binding Proteins. Mol. Cell. Biol. *18* , 6538–6547.

Hendrich, B., Guy, J., Ramsahoye, B., Wilson, V.A., and Bird, A. (2001). Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development. Genes Dev. *15*, 710–723.

Higashi, Y., Maruhashi, M., Nelles, L., Van de Putte, T., Verschueren, K., Miyoshi, T., Yoshimoto, A., Kondoh, H., and Huylebroeck, D. (2002a). Generation of the floxed allele of the SIP1 (Smad-interacting protein 1) gene for Cre-mediated conditional knockout in the mouse. Genesis *32*, 82–84.

Higashi, Y., Maruhashi, M., Nelles, L., Putte, T. Van De, Verschueren, K., Miyoshi, T., Yoshimoto, A., Kondoh, H., and Huylebroeck, D. (2002b). Generation of the Floxed Allele of the SIP1 (Smad-Interacting Protein 1) Gene for Cre-Mediated Conditional Knockout in the Mouse. *84*, 82–84.

Ho, L., Ronan, J.L., Wu, J., Staahl, B.T., Chen, L., Kuo, A., Lessard, J., Nesvizhskii, A.I., Ranish, J., and Crabtree, G.R. (2009). An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. Proc. Natl. Acad. Sci. *106* , 5181–5186.

Hoch, R. V, Rubenstein, J.L.R., and Pleasure, S. (2009). Genes and signaling events that establish regional patterning of the mammalian forebrain. Semin. Cell Dev. Biol. *20*, 378–386.

Hoerder-Suabedissen, A., and Molnar, Z. (2015). Development, evolution and pathology of neocortical subplate neurons. Nat Rev Neurosci *16*, 133–146.

Houbaviy, H.B., Murray, M.F., and Sharp, P.A. (2003). Embryonic Stem Cell-Specific MicroRNAs. Dev. Cell *5*, 351–358.

Hu, X., Zhang, L., Mao, S.-Q., Li, Z., Chen, J., Zhang, R.-R., Wu, H.-P., Gao, J., Guo, F., Liu, W., et al. (2014). Tet and TDG mediate DNA demethylation essential for mesenchymal-to-epithelial transition in somatic cell reprogramming. Cell Stem Cell *14*, 512–522.

Huo, J.S., and Zambidis, E.T. (2013). Pivots of pluripotency: the roles of non-coding RNA in regulating embryonic and induced pluripotent stem cells. Biochim. Biophys. Acta *1830*, 2385–2394.

Ito, S., D'Alessio, A.C., Taranova, O. V, Hong, K., Sowers, L.C., and Zhang, Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature *466*, 1129–1133.

Ivanova, N., Dobrin, R., Lu, R., Kotenko, I., Levorse, J., DeCoste, C., Schafer, X., Lun, Y., and Lemischka, I.R. (2006). Dissecting self-renewal in stem cells with RNA interference. Nature *442*, 533–538.

Iwafuchi-Doi, M., Matsuda, K., Murakami, K., Niwa, H., Tesar, P.J., Aruga, J., Matsuo, I., and Kondoh, H. (2012). Transcriptional regulatory networks in epiblast cells and during anterior neural plate development as modeled in epiblast stem cells. Development *139*, 4675–4675.

Jackson, M., Krassowska, A., Gilbert, N., Chevassut, T., Forrester, L., Ansell, J., and Ramsahoye, B. (2004). Severe Global DNA Hypomethylation Blocks Differentiation and Induces Histone Hyperacetylation in Embryonic Stem Cells. Mol. Cell. Biol. *24* , 8862–8871.

Jamaladdin, S., Kelly, R.D.W., O'Regan, L., Dovey, O.M., Hodson, G.E., Millard, C.J., Portolano, N., Fry, A.M., Schwabe, J.W.R., and Cowley, S.M. (2014). Histone deacetylase (HDAC) 1 and 2 are essential for accurate cell division and the pluripotency of embryonic stem cells. Proc. Natl. Acad. Sci. U. S. A. *111*, 9840–9845.

Jang, H., Kim, T.W., Yoon, S., Choi, S.-Y., Kang, T.-W., Kim, S.-Y., Kwon, Y.-W., Cho, E.-J., and Youn, H.-D. (2012). O-GlcNAc Regulates Pluripotency and Reprogramming by Directly Acting on Core Components of the Pluripotency Network. Cell Stem Cell *11*, 62–74.

Jedrusik, A., Parfitt, D.-E., Guo, G., Skamagki, M., Grabarek, J.B., Johnson, M.H., Robson, P., and Zernicka-Goetz, M. (2008). Role of Cdx2 and cell polarity in cell allocation and specification of trophectoderm and inner cell mass in the mouse embryo. Genes Dev. *22* , 2692–2706.

Jeong, M., Sun, D., Luo, M., Huang, Y., Challen, G.A., Rodriguez, B., Zhang, X., Chavez, L., Wang, H., Hannah, R., et al. (2014a). Large conserved domains of low DNA methylation maintained by Dnmt3a. Nat. Genet. *46*, 17–23.

Jeong, M., Sun, D., Luo, M., Huang, Y., Challen, G.A., Rodriguez, B., Zhang, X., Chavez, L., Wang, H., Hannah, R., et al. (2014b). Large conserved domains of low DNA methylation maintained by Dnmt3a. Nat. Genet. *46*, 17–23.

Jeub, M., Emrich, M., Pradier, B., Taha, O., Gailus-Durner, V., Fuchs, H., de Angelis, M.H., Huylebroeck, D., Zimmer, A., Beck, H., et al. (2011). The transcription factor Smad-interacting protein 1 controls pain sensitivity via modulation of DRG neuron excitability. Pain *152*, 2384–2398.

Van der Jeught, M., Taelman, J., Duggal, G., Ghimire, S., Lierman, S., Chuva de Sousa Lopes, S.M., Deforce, D., Deroo, T., De Sutter, P., and Heindryckx, B. (2015). Application Of Small Molecules Favoring Naïve Pluripotency during Human Embryonic Stem Cell Derivation. Cell. Reprogram. *17*, 170–180.

Jia, J., Zheng, X., Hu, G., Cui, K., Zhang, J., Zhang, A., Jiang, H., Lu, B., Yates III, J., Liu, C., et al. (2012). Regulation of Pluripotency and Self- Renewal of ESCs through Epigenetic- Threshold Modulation and mRNA Pruning. Cell *151*, 576–589.

Joo, J.Y., Choi, H.W., Kim, M.J., Zaehres, H., Tapia, N., Stehling, M., Jung, K.S., Do, J.T., and Schöler, H.R. (2014). Establishment of a primed pluripotent epiblast stem cell in FGF4-based conditions. Sci. Rep. *4*, 7477.

Joshi, O., Wang, S.-Y., Kuznetsova, T., Atlasi, Y., Peng, T., Fabre, P.J., Habibi, E., Shaik, J., Saeed, S., Handoko, L., et al. (2015). Dynamic Reorganization of Extremely Long-Range Promoter-Promoter Interactions between Two States of Pluripotency. Cell Stem Cell *17*, 748–757.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S., et al. (2010). Mediator and cohesin connect gene expression and chromatin architecture. Nature *467*, 430–435.

Kahan, B.W., and Ephrussi, B. (1970). Developmental potentialities of clonal in vitro cultures of mouse testicular teratoma. J. Natl. Cancer Inst. *44*, 1015–1036.

Kahlert, C., Lahes, S., Radhakrishnan, P., Dutta, S., Mogler, C., Herpel, E., Brand, K., Steinert, G., Schneider, M., Mollenhauer, M., et al.

(2011). Overexpression of ZEB2 at the Invasion Front of Colorectal Cancer Is an Independent Prognostic Marker and Regulates Tumor Invasion In Vitro . Clin. Cancer Res. *17* , 7654–7663.

Kaji, K., Caballero, I.M.M., MacLeod, R., Nichols, J., Wilson, V.A., and Hendrich, B. (2006). The NuRD component Mbd3 is required for pluripotency of embryonic stem cells. Nat Cell Biol *8*, 285–292.

Kalkan, T., and Smith, A. (2014). Mapping the route from naive pluripotency to lineage specification. Philos. Trans. R. Soc. London B Biol. Sci. *369*.

Kanellopoulou, C., Muljo, S.A., Kung, A.L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D.M., and Rajewsky, K. (2005). Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev. *19* , 489–501.

Kang, M., Piliszek, A., Artus, J., and Hadjantonakis, A.-K. (2013). FGF4 is required for lineage restriction and salt-and-pepper distribution of primitive endoderm factors but not their initial expression in the mouse. Dev. *140* , 267–279.

Kaplan, D.R., and Miller, F.D. (2000). Neurotrophin signal transduction in the nervous system. Curr. Opin. Neurobiol. *10*, 381–391.

Karwacki-Neisius, V., Göke, J., Osorno, R., Halbritter, F., Ng, J.H., Weiße, A.Y., Wong, F.C.K., Gagliardi, A., Mullin, N.P., Festuccia, N., et al. (2013). Reduced Oct4 expression directs a robust pluripotent state with distinct signaling activity and increased enhancer occupancy by Oct4 and Nanog. Cell Stem Cell *12*, 531–545.

Kelly, K.F., Ng, D.Y., Jayakumaran, G., Wood, G.A., Koide, H., and Doble, B.W. (2011). β-Catenin Enhances Oct-4 Activity and Reinforces Pluripotency through a TCF-Independent Mechanism. Cell Stem Cell *8*, 214–227.

Kettenmann, H., Hanisch, U.-K., Noda, M., and Verkhratsky, A. (2011). Physiology of Microglia. Physiol. Rev. *91*, 461–553.

Kidder, B.L., Palmer, S., and Knott, J.G. (2009). SWI/SNF-Brg1 Regulates Self-Renewal and Occupies Core Pluripotency-Related Genes in Embryonic Stem Cells. Stem Cells *27*, 317–328.

Kim, D.-K., Cha, Y., Ahn, H.-J., Kim, G., and Park, K.-S. (2013). Lefty1 and Lefty2 Control the Balance Between Self-Renewal and Pluripotent Differentiation of Mouse Embryonic Stem Cells. Stem Cells Dev. *23*, 457–466.

Kim, D.H., Marinov, G.K., Pepke, S., Singer, Z.S., He, P., Williams, B., Schroth, G.P., Elowitz, M.B., and Wold, B.J. (2015). Single-Cell Transcriptome Analysis Reveals Dynamic Changes in lncRNA Expression during Reprogramming. Cell Stem Cell *16*, 88–101.

Kim, S., Lehtinen, M.K., Sessa, A., Zappaterra, M.W., Cho, S.-H., Gonzalez, D., Boggan, B., Austin, C.A., Wijnholds, J., Gambello, M.J., et al. (2010). The apical complex couples cell fate and cell survival to cerebral cortical development. Neuron *66*, 69–84.

Kim, V.N., Han, J., and Siomi, M.C. (2009). Biogenesis of small RNAs in animals. Nat Rev Mol Cell Biol *10*, 126–139.

Kim, Y.-S., Yi, B.-R., Kim, N.-H., and Choi, K.-C. (2014). Role of the epithelial-mesenchymal transition and its effects on embryonic stem cells. Exp Mol Med *46*, e108.

Klein, A.M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D.A., and Kirschner, M.W. (2015). Droplet Barcoding for Single-Cell Transcriptomics Applied to Embryonic Stem Cells. Cell *161*, 1187–1201.

Kleinsmith, L.J., and Pierce, G.B. (1964). Multipotentiality of Single Embryonal Carcinoma Cells. Cancer Res. *24* , 1544–1551.

Koh, K.P., Yabuuchi, A., Rao, S., Huang, Y., Cunniff, K., Nardone, J., Laiho, A., Tahiliani, M., Sommer, C.A., Mostoslavsky, G., et al. (2011). Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. Cell Stem Cell *8*, 200–213.

Kohli, R.M., and Zhang, Y. (2013). TET enzymes, TDG and the dynamics of DNA demethylation. Nature *502*, 472–479.

Kolodziejczyk, A.A., Kim, J.K., Tsang, J.C.H., Ilicic, T., Henriksson, J., Natarajan, K.N., Tuck, A.C., Gao, X., Bühler, M., Liu, P., et al. (2015). Single Cell RNA-Sequencing of Pluripotent States Unlocks Modular Transcriptional Variation. Cell Stem Cell *17*, 471–485.

Korpal, M., Lee, E.S., Hu, G., and Kang, Y. (2008). The miR-200 Family Inhibits Epithelial-Mesenchymal Transition and Cancer Cell Migration by Direct Targeting of E-cadherin Transcriptional Repressors ZEB1 and ZEB2. J. Biol. Chem. *283* , 14910–14914.

Kunath, T., Saba-El-Leil, M.K., Almousailleakh, M., Wray, J., Meloche, S., and Smith, A. (2007). FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. Development *134*, 2895–2902.

de la Serna, I.L., Ohkawa, Y., and Imbalzano, A.N. (2006). Chromatin remodelling in mammalian differentiation: lessons from ATP-dependent remodellers. Nat Rev Genet *7*, 461–473.

Lamouille, S., Xu, J., and Derynck, R. (2014). Molecular mechanisms of epithelial–mesenchymal transition. Nat. Rev. Mol. Cell Biol. *15*, 178–196.

Lancaster, M.A., Renner, M., Martin, C.-A., Wenzel, D., Bicknell, L.S., Hurles, M.E., Homfray, T., Penninger, J.M., Jackson, A.P., and Knoblich, J.A. (2013). Cerebral organoids model human brain development and microcephaly. Nature *501*, 373–379.

Landry, J., Sharov, A.A., Piao, Y., Sharova, L. V, Xiao, H., Southon, E., Matta, J., Tessarollo, L., Zhang, Y.E., Ko, M.S.H., et al. (2008). Essential Role of Chromatin Remodeling Protein Bptf in Early Mouse Embryos and Embryonic Stem Cells. PLoS Genet. *4*, e1000241.

Lanner, F., and Rossant, J. (2010). The role of FGF/Erk signaling in pluripotent cells.

Lee, T.I., Johnstone, S.E., and Young, R.A. (2006). Chromatin immunoprecipitation and microarray-based analysis of protein location. Nat. Protoc. *1*, 729–748.

Leeb, M., and Wutz, A. (2007). Ring1B is crucial for the regulation of developmental control genes and PRC1 proteins but not X inactivation in embryonic cells. J. Cell Biol. *178*, 219–229.

Leeb, M., Dietmann, S., Paramor, M., Niwa, H., and Smith, A. (2014). Genetic exploration of the exit from self-renewal using haploid embryonic stem cells. Cell Stem Cell *14*, 385–393.

Lehtinen, M.K., Zappaterra, M.W., Chen, X., Yang, Y.J., Hill, A.D., Lun, M., Maynard, T., Gonzalez, D., Kim, S., Ye, P., et al. (2011). The cerebrospinal fluid provides a proliferative niche for neural progenitor cells. Neuron *69*, 893–905.

Leitch, H.G., Blair, K., Mansfield, W., Ayetey, H., Humphreys, P., Nichols, J., Surani, M.A., and Smith, A. (2010). Embryonic germ cells from mice and rats exhibit properties consistent with a generic pluripotent ground state. Development *137*, 2279–2287.

Leitch, H.G., McEwen, K.R., Turp, A., Encheva, V., Carroll, T., Grabole, N., Mansfield, W., Nashun, B., Knezovich, J.G., Smith, A., et al.

(2013). Naive pluripotency is associated with global DNA hypomethylation. Nat Struct Mol Biol *20*, 311–316.
Lerchner, W., Latinkic, B. V, Remacle, J.E., Huylebroeck, D., and Smith, J.C. (2000). Region-specific activation of the Xenopus brachyury promoter involves active repression in ectoderm and endoderm: a study using transgenic frog embryos. Development *127*, 2729–2739.
Levasseur, D.N., Wang, J., Dorschner, M.O., Stamatoyannopoulos, J.A., and Orkin, S.H. (2008). Oct4 dependence of chromatin structure within the extended Nanog locus in ES cells. Genes Dev. *22*, 575–580.
Li, B., Carey, M., and Workman, J.L. (2007a). The role of chromatin during transcription. Cell *128*, 707–719.
Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell *69*, 915–926.
Li, J.-Y., Pu, M.-T., Hirasawa, R., Li, B.-Z., Huang, Y.-N., Zeng, R., Jing, N.-H., Chen, T., Li, E., Sasaki, H., et al. (2007b). Synergistic function of DNA methyltransferases Dnmt3a and Dnmt3b in the methylation of Oct4 and Nanog. Mol. Cell. Biol. *27*, 8748–8759.
Li, X., Li, L., Pandey, R., Byun, J.S., Gardner, K., Qin, Z., and Dou, Y. (2012a). The histone acetyltransferase MOF is a key regulator of the embryonic stem cell core transcriptional network. Cell Stem Cell *11*, 163–178.
Li, X., Newbern, J.M., Wu, Y., Morgan-Smith, M., Zhong, J., Charron, J., and Snider, W.D. (2012b). MEK Is a Key Regulator of Gliogenesis in the Developing Brain. Neuron *75*, 1035–1050.
Li, Y., McClintick, J., Zhong, L., Edenberg, H.J., Yoder, M.C., and Chan, R.J. (2004). Murine embryonic stem cell differentiation is promoted by SOCS-3 and inhibited by the zinc finger transcription factor Klf4. Blood *105*, 635–637.
Li, Z., Cai, X., Cai, C.-L., Wang, J., Zhang, W., Petersen, B.E., Yang, F.-C., and Xu, M. (2011). Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. Blood *118*, 4509–4518.
Liao, J., Karnik, R., Gu, H., Ziller, M.J., Clement, K., Tsankov, A.M., Akopian, V., Gifford, C.A., Donaghey, J., Galonska, C., et al. (2015). Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. Nat Genet *47*, 469–478.
Lin, R., and Iacovitti, L. (2015). Classic and novel stem cell niches in brain homeostasis and repair. Brain Res.
Lin, Y., Li, X.-Y., Willis, A.L., Liu, C., Chen, G., and Weiss, S.J. (2014). Snail1-dependent control of embryonic stem cell pluripotency and lineage commitment. Nat Commun *5*.
Liu, W., Ye, P., O'Kusky, J.R., and D'Ercole, A.J. (2009). Type 1 insulin-like growth factor receptor signaling is essential for the development of the hippocampal formation and dentate gyrus. J. Neurosci. Res. *87*, 2821–2832.
Loh, Y.-H., Wu, Q., Chew, J.-L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., et al. (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat. Genet. *38*, 431–440.
Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. bioRxiv.
Maden, M. (2002). Retinoid signalling in the development of the central nervous system. Nat Rev Neurosci *3*, 843–853.
Mak, W., Nesterova, T.B., de Napoles, M., Appanah, R., Yamanaka, S., Otte, A.P., and Brockdorff, N. (2004). Reactivation of the Paternal X Chromosome in Early Mouse Embryos. Sci. *303* , 666–669.
Marks, H., Kalkan, T., Menafra, R., Denissov, S., Jones, K., Hofemeister, H., Nichols, J., Kranz, A., Francis Stewart, A., Smith, A., et al. (2014). The Transcriptional and Epigenomic Foundations of Ground State Pluripotency. Cell *149*, 590–604.
Marson, A., Levine, S.S., Cole, M.F., Frampton, G.M., Brambrink, T., Johnstone, S., Guenther, M.G., Johnston, W.K., Wernig, M., Newman, J., et al. (2008). Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell *134*, 521–533.
Martello, G., Sugimoto, T., Diamanti, E., Joshi, A., Hannah, R., Ohtsuka, S., Göttgens, B., Niwa, H., and Smith, A. (2012). Esrrb is a pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell self-renewal. Cell Stem Cell *11*, 491–504.
Martello, G., Bertone, P., and Smith, A. (2013). Identification of the missing pluripotency mediator downstream of leukaemia inhibitory factor. EMBO J. *32*, 2561–2574.
Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells Developmental Biology : *78*, 7634–7638.
Martinez, S., Crossley, P.H., Cobos, I., Rubenstein, J.L., and Martin, G.R. (1999). FGF8 induces formation of an ectopic isthmic organizer and isthmocerebellar development via a repressive effect on Otx2 expression. Dev. *126* , 1189–1200.
Maruhashi, M., Van De Putte, T., Huylebroeck, D., Kondoh, H., and Higashi, Y. (2005). Involvement of SIP1 in positioning of somite boundaries in the mouse embryo. Dev. Dyn. *234*, 332–338.
Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A.A., et al. (2007). Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. Nat Cell Biol *9*, 625–635.
Matsui, Y., Zsebo, K., and Hogan, B.L.M. (1992). Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. Cell *70*, 841–847.
McConnell, S.K., and Kaznowski, C.E. (1991). No Title. Science (80-.). *254*, 282.
McKinsey, G.L.L., Lindtner, S., Trzcinski, B., Visel, A., Pennacchio, L.A.A., Huylebroeck, D., Higashi, Y., and Rubenstein, J.L.R.L.R. (2013). Dlx1&2-Dependent Expression of Zfhx1b (Sip1, Zeb2) Regulates the Fate Switch between Cortical and Striatal Interneurons. Neuron *77*, 83–98.
Medvinsky, A., and Dzierzak, E. (1996). Definitive Hematopoiesis Is Autonomously Initiated by the AGM Region. Cell *86*, 897–906.
Meersseman, G., Verschueren, K., Nelles, L., Blumenstock, C., Kraft, H., Wuytens, G., Remacle, J., Kozak, C.A., Tylzanowski, P., Niehrs, C., et al. (1997). The C-terminal domain of Mad-like signal transducers is sufficient for biological activity in the Xenopus embryo and transcriptional activation. Mech. Dev. *61*, 127–140.
Métin, C., Baudoin, J.-P., Rakić, S., and Parnavelas, J.G. (2006). Cell and molecular mechanisms involved in the migration of cortical interneurons. Eur. J. Neurosci. *23*, 894–900.
Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.-K., Koche, R.P., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553–560.

Mikkola, H.K.A., and Orkin, S.H. (2006). The journey of developing hematopoietic stem cells. Dev. *133* , 3733–3744.
Miquelajauregui, A., Van de Putte, T., Polyakov, A., Nityanandam, A., Boppana, S., Seuntjens, E., Karabinos, A., Higashi, Y., Huylebroeck, D., and Tarabykin, V. (2007). Smad-interacting protein-1 (Zfhx1b) acts upstream of Wnt signaling in the mouse hippocampus and controls its formation. Proc. Natl. Acad. Sci. *104* , 12919–12924.
Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell *113*, 631–642.
Miyoshi, A., Kitajima, Y., Sumi, K., Sato, K., Hagiwara, A., Koga, Y., and Miyazaki, K. (2004). Snail and SIP1 increase cancer invasion by upregulating MMP family in hepatocellular carcinoma cells. Br J Cancer *90*, 1265–1273.
Molofsky, A. V, Krenick, R., Ullian, E., Tsai, H., Deneen, B., Richardson, W.D., Barres, B.A., and Rowitch, D.H. (2012). Astrocytes and disease: a neurodevelopmental perspective. Genes Dev. *26* , 891–907.
Molyneaux, B.J., Arlotta, P., Menezes, J.R.L., and Macklis, J.D. (2007). Neuronal subtype specification in the cerebral cortex. Nat Rev Neurosci *8*, 427–437.
Montgomery, R.L., Davis, C.A., Potthoff, M.J., Haberland, M., Fielitz, J., Qi, X., Hill, J.A., Richardson, J.A., and Olson, E.N. (2007). Histone deacetylases 1 and 2 redundantly regulate cardiac morphogenesis, growth, and contractility. Genes Dev. *21*, 1790–1802.
Morikawa, M., Koinuma, D., Mizutani, A., Kawasaki, N., Holmborn, K., Sundqvist, A., Tsutsumi, S., Watabe, T., Aburatani, H., Heldin, C.-H., et al. (2016). BMP Sustains Embryonic Stem Cell Self-Renewal through Distinct Functions of Different Krüppel-like Factors. Stem Cell Reports *6*, 64–73.
Morris, S.A., Graham, S.J.L., Jedrusik, A., and Zernicka-Goetz, M. (2013). The differential response to Fgf signalling in cells internalized at different times influences lineage segregation in preimplantation mouse embryos. Open Biol. *3*.
Moshkin, Y.M., Mohrmann, L., van Ijcken, W.F.J., and Verrijzer, C.P. (2007). Functional Differentiation of SWI/SNF Remodelers in Transcription and Cell Cycle Control. Mol. Cell. Biol. *27* , 651–661.
Mostoslavsky, R., Chua, K.F., Lombard, D.B., Pang, W.W., Fischer, M.R., Gellon, L., Liu, P., Mostoslavsky, G., Franco, S., Murphy, M.M., et al. (2006). Genomic instability and aging-like phenotype in the absence of mammalian SIRT6. Cell *124*, 315–329.
Mowat, D.R., Croaker, G.D., Cass, D.T., Kerr, B.A., Chaitow, J., Adès, L.C., Chia, N.L., and Wilson, M.J. (1998). Hirschsprung disease, microcephaly, mental retardation, and characteristic facial features: delineation of a new syndrome and identification of a locus at chromosome 2q22-q23. J. Med. Genet. *35*, 617–623.
Müller, A.M., Medvinsky, A., Strouboulis, J., Grosveld, F., and Dzierzakt, E. (1994). Development of hematopoietic stem cell activity in the mouse embryo. Immunity *1*, 291–301.
Nakaki, F., and Saitou, M. (2014). PRDM14: a unique regulator for pluripotency and epigenetic reprogramming. Trends Biochem. Sci. *39*, 289–298.
Nichols, J., and Smith, A. (2009). Naive and primed pluripotent states. Cell Stem Cell *4*, 487–492.
Nichols, J., and Smith, A. (2012). Pluripotency in the Embryo and in Culture. Cold Spring Harb. Perspect. Biol. *4* .
Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Schöler, H., and Smith, A. (1998). Formation of Pluripotent Stem Cells in the Mammalian Embryo Depends on the POU Transcription Factor Oct4. Cell *95*, 379–391.
Nichols, J., Silva, J., Roode, M., and Smith, A. (2009). Suppression of Erk signalling promotes ground state pluripotency in the mouse embryo. Development *136*, 3215–3222.
Nishioka, N., Inoue, K., Adachi, K., Kiyonari, H., Ota, M., Ralston, A., Yabuta, N., Hirahara, S., Stephenson, R.O., Ogonuki, N., et al. (2009). The Hippo signaling pathway components Lats and Yap pattern Tead4 activity to distinguish mouse trophectoderm from inner cell mass. Dev. Cell *16*, 398–410.
Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet *24*, 372–376.
O'Carroll, D., Erhardt, S., Pagani, M., Barton, S.C., Surani, M.A., and Jenuwein, T. (2001). The polycomb-group gene Ezh2 is required for early mouse development. Mol. Cell. Biol. *21*, 4330–4336.
Ogawa, K., Nishinakamura, R., Iwamatsu, Y., Shimosato, D., and Niwa, H. (2006). Synergistic action of Wnt and LIF in maintaining pluripotency of mouse ES cells. Biochem. Biophys. Res. Commun. *343*, 159–166.
OHTSUKA, M., SOUMIYA, H., HANAI, M., FURUKAWA, S., and FUKUMITSU, H. (2013). Neurotrophin-3 influences the number and the laminar fate of cortical progenitors in the developing cerebral cortex of mice through the MEK/ERK1/2 signaling pathway. Biomed. Res. *34*, 231–239.
Okada, Y., Shimazaki, T., Sobue, G., and Okano, H. (2004). Retinoic-acid-concentration-dependent acquisition of neural cell identity during in vitro differentiation of mouse embryonic stem cells. Dev. Biol. *275*, 124–142.
Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development. Cell *99*, 247–257.
Orkin, S.H., and Zon, L.I. (2008). Hematopoiesis: an evolving paradigm for stem cell biology. Cell *132*, 631–644.
Osorno, R., Tsakiridis, A., Wong, F., Cambray, N., Economou, C., Wilkie, R., Blin, G., Scotting, P.J., Chambers, I., and Wilson, V. (2012). The developmental dismantling of pluripotency is reversed by ectopic Oct4 expression. Dev. *139* , 2288–2298.
Papatsenko, D., Darr, H., Kulakovskiy, I.V., Waghray, A., Makeev, V.J., MacArthur, B.D., and Lemischka, I.R. (2015). Single-Cell Analyses of ESCs Reveal Alternative Pluripotent Cell States and Molecular Mechanisms that Control Self-Renewal. Stem Cell Reports *5*, 207–220.
Pardo, M., Lang, B., Yu, L., Prosser, H., Bradley, A., Babu, M.M., and Choudhary, J. (2010). An expanded Oct4 interaction network: implications for stem cell biology, development, and disease. Cell Stem Cell *6*, 382–395.
Park, K.-S., Cha, Y., Kim, C.-H., Ahn, H.-J., Kim, D., Ko, S., Kim, K.-H., Chang, M.-Y., Ko, J.-H., Noh, Y.-S., et al. (2013). Transcription elongation factor Tcea3 regulates the pluripotent differentiation potential of mouse embryonic stem cells, via the Lefty1-Nodal-Smad2 pathway. Stem Cells *31*, 282–292.
Park, S.-M., Gaur, A.B., Lengyel, E., and Peter, M.E. (2008). The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. Genes Dev. *22*, 894–907.

Parthasarathy, S., Srivatsa, S., Nityanandam, A., and Tarabykin, V. (2014). Ntf3 acts downstream of Sip1 in cortical postmitotic neurons to control progenitor cell fate through feedback signaling. Dev. *141* , 3324–3330.

Pasini, D., Bracken, A.P., Jensen, M.R., Denchi, E.L., and Helin, K. (2004). Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. EMBO J. *23*, 4061–4071.

Pasini, D., Bracken, A.P., Hansen, J.B., Capillo, M., and Helin, K. (2007). The Polycomb Group Protein Suz12 Is Required for Embryonic Stem Cell Differentiation. Mol. Cell. Biol. *27* , 3769–3779.

Pieters, T., Haenebalcke, L., Hochepied, T., D'Hont, J., Haigh, J.J., van Roy, F., van Hengel, J., D'Hont, J., Haigh, J.J., van Roy, F., et al. (2012). Efficient and User-Friendly Pluripotin-based Derivation of Mouse Embryonic Stem Cells. Stem Cell Rev. Reports *8*, 768–778.

Pieters&VanRoy (2014). Role of cell-cell adhesion complexes in embryonic stem cell biology. J. CELL Sci. 127(12), 2603–2613.

Postigo, A.A., Depp, J.L., Taylor, J.J., and Kroll, K.L. (2003). Regulation of Smad signaling through a differential recruitment of coactivators and corepressors by ZEB proteins. EMBO J. *22*, 2453–2462.

Pourfarzad, F., Aghajanirefah, A., de Boer, E., Ten Have, S., Bryn van Dijk, T., Kheradmandkia, S., Stadhouders, R., Thongjuea, S., Soler, E., Gillemans, N., et al. (2013). Locus-specific proteomics by TChP: targeted chromatin purification. Cell Rep. *4*, 589–600.

Powell, E.M., Campbell, D.B., Stanwood, G.D., Davis, C., Noebels, J.L., and Levitt, P. (2003). Genetic Disruption of Cortical Interneuron Development Causes Region- and GABA Cell Type-Specific Deficits, Epilepsy, and Behavioral Dysfunction. J. Neurosci. *23* , 622–631.

Van de Putte, T., Maruhashi, M., Francis, A., Nelles, L., Kondoh, H., Huylebroeck, D., and Higashi, Y. (2003). Mice lacking ZFHX1B, the gene that codes for Smad-interacting protein-1, reveal a role for multiple neural crest cell defects in the etiology of Hirschsprung disease-mental retardation syndrome. Am. J. Hum. Genet. *72*, 465–470.

Van de Putte, T., Francis, A., Nelles, L., van Grunsven, L.A., and Huylebroeck, D. (2007). Neural crest-specific removal of Zfhx1b in mouse leads to a wide range of neurocristopathies reminiscent of Mowat-Wilson syndrome. Hum. Mol. Genet. *16*, 1423–1436.

Qian, X., Kim, J.K., Tong, W., Villa-Diaz, L.G., and Krebsbach, P.H. (2015). DPPA5 Supports Pluripotency and Reprogramming by Regulating NANOG Turnover. Stem Cells n/a – n/a.

Rakic, P. (2009). Evolution of the neocortex: Perspective from developmental biology. Nat. Rev. Neurosci. *10*, 724–735.

Remacle, J.E., Kraft, H., Lerchner, W., Wuytens, G., Collart, C., Verschueren, K., Smith, J.C., and Huylebroeck, D. (1999). New mode of DNA binding of multi-zinc finger transcription factors: deltaEF1 family members bind with two hands to two target sites. EMBO J. *18*, 5073–5084.

Reynolds, N., Latos, P., Hynes-Allen, A., Loos, R., Leaford, D., O'Shaughnessy, A., Mosaku, O., Signolet, J., Brennecke, P., Kalkan, T., et al. (2012). NuRD suppresses pluripotency gene expression to promote transcriptional heterogeneity and lineage commitment. Cell Stem Cell *10*, 583–594.

Ribeiro Xavier, A.L., Kress, B.T., Goldman, S.A., Lacerda de Menezes, J.R., and Nedergaard, M. (2015). A Distinct Population of Microglia Supports Adult Neurogenesis in the Subventricular Zone. J. Neurosci. *35* , 11848–11861.

Ringrose, L., and Paro, R. (2004). Epigenetic Regulation of Cellular Memory by the Polycomb and Trithorax Group Proteins. Annu. Rev. Genet. *38*, 413–443.

Rosivatz, E., Becker, I., Specht, K., Fricke, E., Luber, B., Busch, R., Höfler, H., and Becker, K.-F. (2015). Differential Expression of the Epithelial-Mesenchymal Transition Regulators Snail, SIP1, and Twist in Gastric Cancer. Am. J. Pathol. *161*, 1881–1891.

Rossant, J., Chazaud, C., and Yamanaka, Y. (2003). Lineage allocation and asymmetries in the early mouse embryo. Philos. Trans. R. Soc. London B Biol. Sci. *358*, 1341–1349.

Roy, A., Francius, C., Rousso, D.L., Seuntjens, E., Debruyn, J., Luxenhofer, G., Huber, A.B., Huylebroeck, D., Novitch, B.G., and Clotman, F. (2012). Onecut transcription factors act upstream of Isl1 to regulate spinal motoneuron diversification. Development *139*, 3109–3119.

Sandell, L.L., Sanderson, B.W., Moiseyev, G., Johnson, T., Mushegian, A., Young, K., Rey, J.-P., Ma, J., Staehling-Hampton, K., and Trainor, P.A. (2007). RDH10 is essential for synthesis of embryonic retinoic acid and is required for limb, craniofacial, and organ development. Genes Dev. *21* , 1113–1124.

Sato, N., Meijer, L., Skaltsounis, L., Greengard, P., and Brivanlou, A.H. (2004). Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. Nat. Med. *10*, 55–63.

Scardigli, R., Bäumer, N., Gruss, P., Guillemot, F., and Le Roux, I. (2003). Direct and concentration-dependent regulation of the proneural gene Neurogenin2 by Pax6. Dev. *130* , 3269–3281.

Schoenfelder, S., Sugar, R., Dimond, A., Javierre, B.-M., Armstrong, H., Mifsud, B., Dimitrova, E., Matheson, L., Tavares-Cadete, F., Furlan-Magaril, M., et al. (2015). Polycomb repressive complex PRC1 spatially constrains the mouse embryonic stem cell genome. Nat Genet *47*, 1179–1186.

Segklia, A., Seuntjens, E., Elkouris, M., Tsalavos, S., Stappers, E., Mitsiadis, T.A., Huylebroeck, D., Remboutsika, E., and Graf, D. (2012). Bmp7 regulates the survival, proliferation, and neurogenic properties of neural progenitor cells during corticogenesis in the mouse. PLoS One *7*, e34088.

Sekido, R., Takagi, T., Okanami, M., Moribe, H., Yamamura, M., Higashi, Y., and Kondoh, H. (1996). Organization of the gene encoding transcriptional repressor δEF1 and cross-species conservation of its domains. Gene *173*, 227–232.

Seuntjens, E., Nityanandam, A., Miquelajauregui, A., Debruyn, J., Stryjewska, A., Goebbels, S., Nave, K.-A., Huylebroeck, D., and Tarabykin, V. (2009). Sip1 regulates sequential fate decisions by feedback signaling from postmitotic neurons to progenitors. Nat. Neurosci. *12*, 1373–1380.

Shafi, R., Iyer, S.P.N., Ellies, L.G., O'Donnell, N., Marek, K.W., Chui, D., Hart, G.W., and Marth, J.D. (2000). The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny. Proc. Natl. Acad. Sci. *97* , 5735–5739.

Shen, Q., Wang, Y., Dimos, J.T., Fasano, C.A., Phoenix, T.N., Lemischka, I.R., Ivanova, N.B., Stifani, S., Morrisey, E.E., and Temple, S. (2006). The timing of cortical neurogenesis is encoded within lineages of individual progenitor cells. Nat Neurosci *9*, 743–751.

Shen, X., Liu, Y., Hsu, Y.-J., Fujiwara, Y., Kim, J., Mao, X., Yuan, G.-C., and Orkin, S.H. (2008). EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. Mol. Cell *32*, 491–502.

Shi, F.-T., Kim, H., Lu, W., He, Q., Liu, D., Goodell, M.A., Wan, M., and Songyang, Z. (2013). Ten-Eleven Translocation 1 (Tet1) Is Regulated by O-Linked N-Acetylglucosamine Transferase (Ogt) for Target Gene Repression in Mouse Embryonic Stem Cells. J. Biol. Chem. *288*, 20776–20784.

Siegenthaler, J.A., Ashique, A.M., Zarbalis, K., Patterson, K.P., Hecht, J.H., Kane, M.A., Folias, A.E., Choe, Y., May, S.R., Kume, T., et al. (2009). Retinoic acid from the meninges regulates cortical neuron generation. Cell *139*, 597–609.

Silva, J., and Smith, A. (2008). Capturing pluripotency. Cell *132*, 532–536.

Silva, J., Nichols, J., Theunissen, T.W., Guo, G., van Oosten, A.L., Barrandon, O., Wray, J., Yamanaka, S., Chambers, I., and Smith, A. (2009). Nanog is the gateway to the pluripotent ground state. Cell *138*, 722–737.

Smallwood, S.A., Lee, H.J., Angermueller, C., Krueger, F., Saadeh, H., Peat, J., Andrews, S.R., Stegle, O., Reik, W., and Kelsey, G. (2014). Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. Nat Meth *11*, 817–820.

Smith, J. (1997). Neurulation: coming to closure. Trends Neurosci. *20*, 510–517.

Smith, A.G., Heath, J.K., Donaldson, D.D., Wong, G.G., Moreau, J., Stahl, M., and Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. Nature *336*, 688–690.

Solter, D., SKREB, N., and DAMJANOV, I. (1970). Extrauterine Growth of Mouse Egg-cylinders results in Malignant Teratoma. Nature *227*, 503–504.

Song, S.J., Poliseno, L., Song, M.S., Ala, U., Webster, K., Ng, C., Beringer, G., Brikbak, N.J., Yuan, X., Cantley, L.C., et al. (2013). MicroRNA-antagonism regulates breast cancer stemness and metastasis via TET-family-dependent chromatin remodeling. Cell *154*, 311–324.

Soriano, P. (1999). Generalized lacZ expression with the ROSA26 Cre reporter strain . Nat Genet *21*, 70–71.

Spassky, N., Merkle, F.T., Flames, N., Tramontin, A.D., García-Verdugo, J.M., and Alvarez-Buylla, A. (2005). Adult Ependymal Cells Are Postmitotic and Are Derived from Radial Glial Cells during Embryogenesis. J. Neurosci. *25* , 10–18.

Stanchina, L., Van de Putte, T., Goossens, M., Huylebroeck, D., and Bondurand, N. (2010). Genetic interaction between Sox10 and Zfhx1b during enteric nervous system development. Dev. Biol. *341*, 416–428.

Stern, C.D. (2005). Neural induction: old problem, new findings, yet more questions. Dev. *132* , 2007–2021.

Stevens, L.C. (1970). The development of transplantable teratocarcinomas from intratesticular grafts of pre- and postimplantation mouse embryos. Dev. Biol. *21*, 364–382.

Su, P., Zhang, J., Zhao, F., Aschner, M., Chen, J., and Luo, W. (2014). The interaction between microglia and neural stem/precursor cells. Brain Res. Bull. *109*, 32–38.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. *102*, 15545–15550.

Sugimoto, M., Kondo, M., Koga, Y., Shiura, H., Ikeda, R., Hirose, M., Ogura, A., Murakami, A., Yoshiki, A., Chuva de Sousa Lopes, S.M., et al. (2015). A Simple and Robust Method for Establishing Homogeneous Mouse Epiblast Stem Cell Lines by Wnt Inhibition. Stem Cell Reports *4*, 744–757.

Sugino, K., Hempel, C.M., Miller, M.N., Hattox, A.M., Shapiro, P., Wu, C., Huang, Z.J., and Nelson, S.B. (2006). Molecular taxonomy of major neuronal classes in the adult mouse forebrain. Nat Neurosci *9*, 99–107.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., et al. (2009). Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. Science *324*, 930–935.

Tai, C.-I., and Ying, Q.-L. (2013). Gbx2, a LIF/Stat3 target, promotes reprogramming to and retention of the pluripotent ground state. J. Cell Sci. *126* , 1093–1098.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663–676.

Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficz, G., Krueger, F., Oxley, D., Santos, F., Clarke, J., Mansfield, W., et al. (2014). Resetting transcription factor control circuitry toward ground-state pluripotency in human. Cell *158*, 1254–1269.

Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and McKay, R.D.G. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature *448*, 196–199.

Theunissen, T.W., Powell, B.E., Wang, H., Mitalipova, M., Faddah, D.A., Reddy, J., Fan, Z.P., Maetzel, D., Ganz, K., Shi, L., et al. (2014). Systematic Identification of Culture Conditions for Induction and Maintenance of Naive Human Pluripotency. Cell Stem Cell *15*, 471–487.

Thomas, T., Dixon, M.P., Kueh, A.J., and Voss, A.K. (2008). Mof (MYST1 or KAT8) Is Essential for Progression of Embryonic Development Past the Blastocyst Stage and Required for Normal Chromatin Architecture . Mol. Cell. Biol. *28*, 5093–5105.

Tolhuis, B., Palstra, R.-J., Splinter, E., Grosveld, F., and de Laat, W. (2002). Looping and Interaction between Hypersensitive Sites in the Active β-globin Locus. Mol. Cell *10*, 1453–1465.

Torres-Padilla, M.-E., Parfitt, D.-E., Kouzarides, T., and Zernicka-Goetz, M. (2007). Histone arginine methylation regulates pluripotency in the early mouse embryo. Nature *445*, 214–218.

Tosolini, M., and Jouneau, A. (2015). From Naive to Primed Pluripotency: In Vitro Conversion of Mouse Embryonic Stem Cells in Epiblast Stem Cells. (Humana Press), pp. 1–8.

Trapnell, C. (2015). Defining cell types and states with single-cell genomics. Genome Res. *25*, 1491–1498.

Trott, J., and Martinez Arias, A. (2013). Single cell lineage analysis of mouse embryonic stem cells at the exit from pluripotency. Biol. Open *2* , 1049–1056.

Tsumura, A., Hayakawa, T., Kumaki, Y., Takebayashi, S., Sakaue, M., Matsuoka, C., Shimotohno, K., Ishikawa, F., Li, E., Ueda, H.R., et al. (2006). Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. Genes to Cells *11*, 805–814.

Turner, S.A., and Bracken, A.P. (2013). A "complex" issue: deciphering the role of variant PRC1 in ESCs. Cell Stem Cell *12*, 145–146.

Urbán, N., and Guillemot, F. (2014). Neurogenesis in the embryonic and adult brain: same regulators, different roles. Front. Cell. Neurosci. *8*, 396.

Vandewalle, C., Comijn, J., De Craene, B., Vermassen, P., Bruyneel, E., Andersen, H., Tulchinsky, E., Van Roy, F., and Berx, G. (2005). SIP1/ZEB2 induces EMT by repressing genes of different epithelial cell–cell junctions. Nucleic Acids Res. *33* , 6566–6578.

Vella, P., Scelfo, A., Jammula, S., Chiacchiera, F., Williams, K., Cuomo, A., Roberto, A., Christensen, J., Bonaldi, T., Helin, K., et al. (2013). Tet proteins connect the O-linked N-acetylglucosamine transferase Ogt to chromatin in embryonic stem cells. Mol. Cell *49*, 645–656.

Verschueren, K., Remacle, J.E., Collart, C., Kraft, H., Baker, B.S., Tylzanowski, P., Nelles, L., Wuytens, G., Su, M.-T., Bodmer, R., et al. (1999). SIP1, a Novel Zinc Finger/Homeodomain Repressor, Interacts with Smad Proteins and Binds to 5'-CACCT Sequences in Candidate Target Genes. J. Biol. Chem. *274* , 20489–20498.

Verstappen, G., van Grunsven, L.A., Michiels, C., Van de Putte, T., Souopgui, J., Van Damme, J., Bellefroid, E., Vandekerckhove, J., and Huylebroeck, D. (2008). Atypical Mowat–Wilson patient confirms the importance of the novel association between ZFHX1B/SIP1 and NuRD corepressor complex. Hum. Mol. Genet. *17* , 1175–1183.

Vleminckx, K., Vakaet, L., Mareel, M., Fiers, W., and Van Roy, F. (1991). Genetic manipulation of E-cadherin expression by epithelial tumor cells reveals an invasion suppressor role. Cell *66*, 107–119.

Wakamatsu, N., Yamada, Y., Yamada, K., Ono, T., Nomura, N., Taniguchi, H., Kitoh, H., Mutoh, N., Yamanaka, T., Mushiake, K., et al. (2001). Mutations in SIP1, encoding Smad interacting protein-1, cause a form of Hirschsprung disease. Nat Genet *27*, 369–370.

Waldrip, Z.J., Byrum, S.D., Storey, A.J., Gao, J., Byrd, A.K., Mackintosh, S.G., Wahls, W.P., Taverna, S.D., Raney, K.D., and Tackett, A.J. (2014). A CRISPR-based approach for proteomic analysis of a single genomic locus. Epigenetics *9*, 1207–1211.

Wang, G., Guo, X., Hong, W., Liu, Q., Wei, T., Lu, C., Gao, L., Ye, D., Zhou, Y., Chen, J., et al. (2013). Critical regulation of miR-200/ZEB2 pathway in Oct4/Sox2-induced mesenchymal-to-epithelial transition and induced pluripotent stem cell generation. Proc. Natl. Acad. Sci. *110* , 2858–2863.

Wang, Y., Medvid, R., Melton, C., Jaenisch, R., and Blelloch, R. (2007). DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. Nat Genet *39*, 380–385.

Wang, Y., Baskerville, S., Shenoy, A., Babiarz, J.E., Baehner, L., and Blelloch, R. (2008). Embryonic stem cell-specific microRNAs regulate the G1-S transition and promote rapid proliferation. Nat Genet *40*, 1478–1483.

Ware, C.B., Nelson, A.M., Mecham, B., Hesson, J., Zhou, W., Jonlin, E.C., Jimenez-Caliani, A.J., Deng, X., Cavanaugh, C., Cook, S., et al. (2014). Derivation of naïve human embryonic stem cells. Proc. Natl. Acad. Sci. *111* , 4484–4489.

Watanabe, K., Kamiya, D., Nishiyama, A., Katayama, T., Nozaki, S., Kawasaki, H., Watanabe, Y., Mizuseki, K., and Sasai, Y. (2005). Directed differentiation of telencephalic precursors from embryonic stem cells. Nat Neurosci *8*, 288–296.

Wei, T., Chen, W., Wang, X., Zhang, M., Chen, J., Zhu, S., Chen, L., Yang, D., Wang, G., Jia, W., et al. (2015). An HDAC2-TET1 switch at distinct chromatin regions significantly promotes the maturation of pre-iPS to iPS cells. Nucleic Acids Res. .

Weng, Q., Chen, Y., Wang, H., Xu, X., Yang, B., He, Q., Shou, W., Chen, Y., Higashi, Y., van den Berghe, V., et al. (2012). Dual-mode Modulation of Smad Signaling by Smad-interacting Protein Sip1 is Required for Myelination in the CNS. Neuron *73*, 713–728.

Wichterle, H., Gifford, D., and Mazzoni, E. (2013). Mapping Neuronal Diversity One Cell at a Time. Sci. *341* , 726–727.

Williams, K., Christensen, J., Pedersen, M.T., Johansen, J. V, Cloos, P.A.C., Rappsilber, J., and Helin, K. (2011). TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature *473*, 343–348.

Williams, R.L., Hilton, D.J., Pease, S., Willson, T.A., Stewart, C.L., Gearing, D.P., Wagner, E.F., Metcalf, D., Nicola, N.A., and Gough, N.M. (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. Nature *336*, 684–687.

Wilson, S.W., and Rubenstein, J.L.R. (2000). Induction and Dorsoventral Patterning of the Telencephalon. Neuron *28*, 641–651.

Wu, H., D/'Alessio, A.C., Ito, S., Xia, K., Wang, Z., Cui, K., Zhao, K., Eve Sun, Y., and Zhang, Y. (2011). Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. Nature *473*, 389–393.

Wurst, W., and Bally-Cuif, L. (2001). Neural plate patterning: Upstream and downstream of the isthmic organizer . Nat Rev Neurosci *2*, 99–108.

Xu, Q., Cobos, I., De La Cruz, E., Rubenstein, J.L., and Anderson, S.A. (2004). Origins of Cortical Interneuron Subtypes. J. Neurosci. *24* , 2612–2622.

Xu, Y., Wu, F., Tan, L., Kong, L., Xiong, L., Deng, J., Barbera, A.J., Zheng, L., Zhang, H., Huang, S., et al. (2011). Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. Mol. Cell *42*, 451–464.

Xue, Y., Wong, J., Moreno, G.T., Young, M.K., Côté, J., and Wang, W. (1998). NURD, a Novel Complex with Both ATP-Dependent Chromatin-Remodeling and Histone Deacetylase Activities. Mol. Cell *2*, 851–861.

Yamaji, M., Ueda, J., Hayashi, K., Ohta, H., Yabuta, Y., Kurimoto, K., Nakato, R., Yamada, Y., Shirahige, K., and Saitou, M. (2013). PRDM14 ensures naive pluripotency through dual regulation of signaling and epigenetic pathways in mouse embryonic stem cells. Cell Stem Cell *12*, 368–382.

Yang, X.-J., and Seto, E. HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention. Oncogene *26*, 5310–5318.

Ye, P., Popken, G.J., Kemper, A., McCarthy, K., Popko, B., and D'Ercole, A.J. (2004). Astrocyte-specific overexpression of insulin-like growth factor-I promotes brain overgrowth and glial fibrillary acidic protein expression. J. Neurosci. Res. *78*, 472–484.

Ye, S., Li, P., Tong, C., and Ying, Q.-L. (2013). Embryonic stem cell self-renewal pathways converge on the transcription factor Tfcp2l1.

EMBO J. *32*, 2548–2560.

Yildirim, O., Li, R., Hung, J.-H., Chen, P.B., Dong, X., Ee, L.-S., Weng, Z., Rando, O.J., and Fazzio, T.G. (2011). Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. Cell *147*, 1498–1510.

Ying, Q.-L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature *453*, 519–523.

Ying, Q.-L.L., Nichols, J., Chambers, I., and Smith, A. (2003). BMP Induction of Id Proteins Suppresses Differentiation and Sustains Embryonic Stem Cell Self-Renewal in Collaboration with STAT3. Cell *115*, 281–292.

Yoshida, K., Chambers, I., Nichols, J., Smith, A., Saito, M., Yasukawa, K., Shoyab, M., Taga, T., and Kishimoto, T. (1994). Maintenance of the pluripotential phenotype of embryonic stem cells through direct activation of gp130 signalling pathways. Mech. Dev. *45*, 163–171.

Yoshimoto, A., Saigou, Y., Higashi, Y., and Kondoh, H. (2005). Regulation of ocular lens development by Smad-interacting protein 1 involving Foxe3 activation. Development *132*, 4437–4448.

Zarbalis, K., Siegenthaler, J.A., Choe, Y., May, S.R., Peterson, A.S., and Pleasure, S.J. (2007). Cortical dysplasia and skull defects in mice with a Foxc1 allele reveal the role of meningeal differentiation in regulating cortical development. Proc. Natl. Acad. Sci. *104* , 14002–14007.

Zeidan, Q., and Hart, G.W. (2010). The intersections between O-GlcNAcylation and phosphorylation: implications for multiple signaling pathways. J. Cell Sci. *123*, 13–22.

Zhang, J., Tam, W.-L., Tong, G.Q., Wu, Q., Chan, H.-Y., Soh, B.-S., Lou, Y., Yang, J., Ma, Y., Chai, L., et al. (2006). Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. Nat Cell Biol *8*, 1114–1123.

Zhang, P., Andrianakos, R., Yang, Y., Liu, C., and Lu, W. (2010). Kruppel-like Factor 4 (Klf4) Prevents Embryonic Stem (ES) Cell Differentiation by Regulating Nanog Gene Expression. J. Biol. Chem. *285*, 9180–9189.

Zhang, Y., Ng, H.-H., Erdjument-Bromage, H., Tempst, P., Bird, A., and Reinberg, D. (1999). Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. Genes Dev. *13*, 1924–1935.

Zhao, S., Hung, F.-C., Colvin, J.S., White, A., Dai, W., Lovicu, F.J., Ornitz, D.M., and Overbeek, P.A. (2001). Patterning the optic neuroepithelium by FGF signaling and Ras activation. Dev. *128* , 5051–5060.

Zhou, Q., Chipperfield, H., Melton, D.A., and Wong, W.H. (2007). A gene regulatory network in mouse embryonic stem cells. Proc. Natl. Acad. Sci. *104* , 16438–16443.

Zweier, C., Thiel, C.T., Dufke, A., Crow, Y.J., Meinecke, P., Suri, M., Ala-Mello, S., Beemer, F., Bernasconi, S., Bianchi, P., et al. (2005). Clinical and mutational spectrum of Mowat-Wilson syndrome. Eur. J. Med. Genet. *48*, 97–111.

Zwijsen, A., Verschueren, K., and Huylebroeck, D. (2003). New intracellular components of bone morphogenetic protein/Smad signaling cascades. FEBS Lett. *546*, 133–139.

C.V.

AGATA STRYJEWSKA
Slachthuislaan 12 bus O 05
3000 Leuven, Belgium
mobile: +32 479 683 996
e-mail: agata-stryjewska@go2.pl
skype: agata.stryjewska

PROFESSIONAL EXPERIENCE:

2008-present	KU Leuven, Faculty of Medicine, Department of Development and Regeneration, Group Embryo and Stem Cells, Laboratory of Molecular Biology (Celgen): PhD student PhD thesis: ***"The role of Zeb2 in cell fate decisions during development"***
03/2007- 12/2007	KU Leuven Biomedical Sciences Group, Department of Human Genetics, Developmental and Molecular Genetics Section, Molecular Biology (Celgen): Laboratory assistant
09/2005-02/2006	Janssen Pharmaceutica and Johnson&Johnson, Internal Medicine Department, Beerse, Belgium Research project within Leonardo da Vinci programme

EDUCATION:

2000-2006	University of Technology, Wroclaw, Poland Master's thesis: "Subunit B of cholera toxin as a small interfering RNA carrier"
02/2005-06/2005	Karel de Grote-Hogeschool**,** Antwerp, Belgium Research project within Socrates-Erasmus programme Thesis: "Enzymatic production of biodiesel: comparison between one-step and a two-step process".

PUBLISHED ARTICLES:

1. Seuntjens E, Nityanandam A, Miquelajauregui A, Debruyn J, **Stryjewska A**, Goebbels S, Nave KA, Huylebroeck D, Tarabykin V. Sip1 regulates sequential fate decisions by feedback signaling from postmitotic neurons to progenitors. ***Nat Neurosci.*** 2009 Nov;12(11):1373-80. Epub 2009 Oct 18

2. Andrea Conidi, Silvia Cazzola, Karen Beets, Kathleen Coddens, Clara Collart, Frederique Cornelis, Luk Cox, Debruyn Joke, Mariya P. Dobreva, Ruben Dries, Camila Esguerra, Annick Francis, Abdelilah Ibrahimi, Roel Kroes, Flore Lesage, Elke Maas, Ivan Moya , Paulo N.G. Pereira, Elke Stappers, **Agata Stryjewska**, Veronique van den Berghe, Liesbeth Vermeire, Griet Verstappen, Eve Seuntjens, Lieve Umans, An Zwijsen, Danny Huylebroeck. Few Smad proteins and many Smad-interacting proteins yield multiple functions and action modes in TGFβ/BMP signaling *in vivo*. ***Cytokine & Growth Factor Rev.* 22**(5): 287-300.

3. Conidi A, van den Berghe V, Leslie K, **Stryjewska A**, Xue H,Chen YG, Seuntjens E, Huylebroeck D.(2013). Four Amino Acids within a Tandem QxVx Repeat in a Predicted Extended α-Helix of the Smad-Binding Domain of Sip1 Are Necessary for Binding to Activated Smad Proteins. ***PLoS ONE*** 8(10): e76733.

4. Omilusik KD, Best JA, Yu B, Goossens S, Weidemann A, Nguyen JV, Seuntjens E, **Stryjewska A,** Zweier C, Roychoudhuri R, Gattinoni L, Bird LM, Higashi Y, Kondoh H, Huylebroeck D, Haigh J, Goldrath AW. (2015). Transcriptional repressor ZEB2 promotes terminal differentiation of CD8+ effector and memory T cell populations during infection. ***J Exp Med.*** 2015 Nov 16;212(12):2027-39

ARTICLES IN REVISION:

1. **Agata Stryjewska**[+], Ruben Dries[+], Tim Pieters, Griet Verstappen, Andrea Conidi, Kathleen Coddens, Annick Francis, Lieve Umans, Wilfred F. J. van IJcken, Geert Berx, Leo A. van Grunsven, Frank Grosveld, Steven Goossens, Jody J. Haigh, Danny Huylebroeck; [+]co-first author. Zeb2 regulates cell fate at the exit from epiblast state in mouse embryonic stem cells.

SUBMITTED ARTICELS:

1. Pieters T, Goosssens S, Haenebalcke L, Andries V, **Stryjewska A**, De Rycke R, Lemeire K, D'Hont J, Van Rockeghem F, Colpaert N, Staes K, Vleeschouwers M, Hochepied T, Huylebroeck D, Berx G, Stemmler M, Wirth D, Haigh JJ, van Hengel J, van Roy F. p120 catenin-mediated stabilization of E-cadherin is essential for proper primitive endoderm formation.

ARTICLES IN PREPARATION:

1. Ruben Dries, **Agata Stryjewska**, Kathleen Coddens, Enrico Glaab, Frank Grosveld, Stein Aerts, Danny Huylebroeck. Understanding the extended TGFbeta pathway as a transcriptional regulatory system in early neural development**.**

2. Ruben Dries, Satoshi Okawa, **Agata Stryjewska**, Kathleen Coddens, Enrico Glaab, Frank Grosveld, Stein Aerts, Antonio del Sol Mesa, Danny Huylebroeck. Dynamic subpopulations and predictive bifurcation analysis in differentiating mouse embryonic stem cells at the single-cell level.

3. Stappers E, van den Berghe V, Dries R, Arredouani S, **Stryjewska A**, Conidi A, Francis A, Bresseleers M, Van Ijcken W, Kessaris N, Götz M, Seuntjens E., Huylebroeck D. Zeb2 contributes to the establishment of the postnatal neurogenic niche.

POSTER AND ORAL PRESENTATIONS:

1. BSCDB meeting, "Cell fate determination during embryogenesis", 30-31/03/2009, Leuven, Belgium, poster presentation: "Bringing mechanism of action studies to mouse ES cells: molecular dissection of the multi-domain transcription factor Sip1 during neural development"
2. Endotrack meeting, 9-11/11/2009, Il Ciocco, Italy, poster presentation: "Bringing mechanism of action studies to mouse ES cells: molecular dissection of the multi-domain transcription factor Sip1 during neural development"
3. Hydra VI Summer School 2010, Stem Cells & Regenerative Medicine, 11-18/09/2010, Hydra, Greece, poster presentation "Bringing mechanism of action studies to mouse ES cells: molecular dissection of the multi-domain transcription factor Sip1 during neural development"
4. SyBoSS meeting "Systems Biology of Stem Cells" 30/01-01/02-2013, Kirchberg, Austria, poster presentation "Comparative transcriptomics for identifying Sip1 dependent genes in neural development"
5. BSCDB meeting, Experimental Models of Human Diseases, 18-19/10/2013, Liege, Belgium, poster presentation: "Comparative transcriptomics for identifying Sip1 dependent genes in neural development"
6. IAP meeting "Signaling in health and disease", 14/12/2013, Gent, Belgium, poster presentation: "*Sip1 in vitro* : Comparative transcriptomics for identifying Sip1 dependent genes in neural development"
7. Symposium, Princess Lilian Foundation, 21/03/2014, oral presentation: "Sip1 *in vitro*: comparative transcriptomics for identifying Sip1 dependent genes in neural development"
8. BSCDB Autumn meeting 2014, Epigenetics and Stem Cell Biology, 6-7/11/2014, Antwerp, Belgium, poster presentation: "Zeb2 controls exit from pluripotency in mouse embryonic stem cells"
9. 5th IUAP network meeting, Rotterdam, The Netherlands, 8-9/12/2014, oral presentation: "Zeb2 regulates exit from pluripotency in mouse embryonic stem cells"
10. Interuniversity Stem Cell Meeting, Leuven, Belgium, 20/04/2015, oral presentation and poster presentation: "Zeb2 regulates exit from pluripotency during neural differentiation of mouse embryonic stem cells"

11. Developmental programing and organ repair, Leuven, Belgium, 23-24/11/2015, poster presentation: “The differentiation transcription factor Zeb2 in mouse ES cells controls pluripotency exit via Tet1-mediated DNA-methylation “